7/6

Contemporary Science

The Basis of Modern Physics

John M Irvine

John M Irvine

The Basis of Modern Physics

12

The Basis of Modern Physics

CONTEMPORARY
SCIENCE
PAPERBACKS

12

J. M. IRVINE B.SC., M.SC., PH.D., A.INST.P.
Lecturer in Theoretical Physics at the University of Manchester

The Basis of Modern Physics

OLIVER & BOYD
EDINBURGH AND LONDON

OLIVER & BOYD LTD
Tweeddale Court Edinburgh 1
39a Welbeck Street London W.1

First published 1967
© 1967 J. M. Irvine

Set in Times New Roman and printed in
Great Britain by Richard Clay (The Chaucer Press) Ltd
Bungay, Suffolk

Preface

This book arose out of a set of extramural lectures given by the author at the University of Manchester throughout the winter of 1965 and the spring of 1966.

Chapter 1 briefly reviews our concepts of classical dynamics and mechanics and the picture of the structure of matter that we obtain from these through the classical kinetic theory. In Chapter 2 we turn our attention to Maxwell's electromagnetic theory and its connection with the nature of light. We include also a short review of some of the basic aspects of wave motion.

We deal next, in Chapter 3, with Einstein's special theory of relativity and its implications for physical measurements.

Chapters 4, 5 and 6 respectively discuss the disagreement between the classical theory and experiments on the interaction of radiation with matter, the attempts to patch up these discrepancies within the *ad hoc* assumptions of the old quantum theory and finally the rise of modern quantum mechanics.

We conclude with a pot-pourri of problems–some partially solved, some very far from solution. In this final chapter the reader is given a glimpse of a few of the fascinating problems confronting the modern physicist.

We have attempted to keep the amount of mathematics used in this book to a minimum. However, in order to keep the size of the book within limits it has been necessary to use some mathematical techniques which may not be familiar to the reader. We have therefore included two appendices, the first on vector analysis and the second on statistics.

Clearly, it is impossible to do justice to the subjects treated here within the framework of such a small volume, and to help the reader we have included a short bibliography where further details may be found.

I should like to thank my wife for her patient reading of the manuscript and also to acknowledge helpful discussions with colleagues both in Manchester and at Cornell University.

It is my wish to dedicate this book to my parents.

J. M. Irvine

Contents

Preface v
1. Classical Mechanics and the Structure of Matter 1
2. Electromagnetism, Light and Waves 20
3. Relativity 40
4. The Interaction of the Electromagnetic Field with Matter 56
5. The Old Quantum Theory 69
6. Quantum Mechanics 81
7. Pot-Pourri 96
Appendix A 110
Appendix B 114
Bibliography 119
Index 121

1. Classical Mechanics and the Structure of Matter

INTRODUCTION

When we talk of modern physics we mean specific branches of physics which have for their theoretical interpretation concepts and ideas developed during the twentieth century. Thus there are many physicists who are working today on what are certainly contemporary problems in physics, but which are not problems in modern physics in the sense that the essential theoretical concepts were, at least in principle, established before the turn of the century.

The bases of classical physics, i.e. Newtonian mechanics, Maxwell's theory of electromagnetism and the statistical mechanics of Boltzmann, were derived from a common-sense interpretation of the observation of phenomena in the macroscopic world. By which is meant that they deal with events which can be observed without a too sophisticated extension of our natural senses. They deal with objects ranging in size from those observable under an optical microscope up to objects the size of the Sun; and with velocities ranging from the state of rest up to the velocity of a planet in its orbit. As man's technical skills have increased and he has attempted to explain observations which are beyond this macroscopic world, flaws have been discovered in the apparent universality of the ideas of classical physics.

The theories of modern physics have been developed to explain the observations beyond the normal limits of the macroscopic world. When we discuss the structure of the atom, an object much too small to be observed through an optical microscope, we require quantum mechanics in place of Newtonian mechanics. When we consider systems moving with velocities approaching that of light, we require to use the theory of relativity in place of our usual concepts of space and time. And when we deal with large numbers of subatomic particles, we require the statistics of Bose and Einstein or Fermi and Dirac in place of the classical statistics of Boltzmann.

It must be stressed that any theory of modern physics must contain, as a limiting aspect of its nature, its counterpart in classical physics in order to accommodate the considerable successes achieved in explaining the macroscopic world.

Many of the objects discussed in modern physics are themselves not directly observable, and their existence and nature can only be inferred from the effects they produce; in particular this is true of the whole world of subatomic particles. This raises the question of what are, and what are not, observable quantities. And if something is not observable, then what physical significance, if any, can we ascribe to it? This becomes particularly important when we discuss a 'gedanken' experiment. This is an imaginary experiment which we might think we could carry out if only we could obtain fine enough instruments, or if some ideal experimental conditions could be satisfied, but which in practice we shall never be able to undertake. The theories of modern physics for the first time set limits for the 'in principle' result of a measurement in such an experiment. We must therefore ensure that when discussing gedanken experiments the questions that we ask do not violate some fundamental statement of principle.

The central concept of physics is that of measurement. The study of theoretical physics is concerned with establishing relationships between the results of different measurements; these relationships are what we know as the laws of physics. However, we must be careful not to extrapolate unquestioningly these laws to realms outside the region of measurement for which they have been established. In all cases these laws are in some sense only an approximation valid within the accuracy and realm of the measurements used to establish them. We may, if we wish, believe that there is some ultimate law and that as our experiments increase in accuracy and

breadth we shall approach this by successive approximations. However, a discussion of whether there is, or is not, some all-embracing truth is not part of the study of physics but belongs rather to the realm of metaphysics.

Before we begin an examination of the theories of modern physics, let us briefly examine a few of the ideas of classical physics.

CLASSICAL MECHANICS

Probably the simplest measurement we can make is one of length. To measure the width of this page, for instance, one would simply place a ruler along it and read off on a scale a number which is related in a very specific way to some standard unit of length, either the standard yard or the standard metre. However, often a simple measurement of length is not a very useful number on its own. Suppose you wish to travel from Edinburgh to London; it is certainly useful to know that the distance between the two cities is 400 miles, but before setting out on such a journey you would also require to know in which direction to travel. A quantity which can be completely described by a magnitude we shall call a scalar. A quantity which requires for its complete specification both a magnitude and a direction we shall call a vector.

In order to describe how to get from point A to point B in Fig. 1, one of the simplest procedures would be to travel a units of length parallel to the bottom edge of the page and then b units of length parallel to the side of the book. This would take us uniquely from A to B.

Obviously any point in the page could be uniquely specified relative to some fixed point O, called the origin, by giving two numbers (x, y) representing the distance which we would require to travel from O parallel to the bottom of the page and then parallel to the side of the page in order to reach the point. Thus the point A is represented by the two numbers (x_1, y_1) and the point B by the numbers (x_2, y_2). The vector AB is then given by the two numbers a and b equal respectively to $(x_2 - x_1)$ and $(y_2 - y_1)$. These two numbers are called the components of the vector, and the numbers (x, y) specifying a point are called the co-ordinates of the point. The origin together with the directions along which the co-ordinates are measured are said to form a frame of reference.

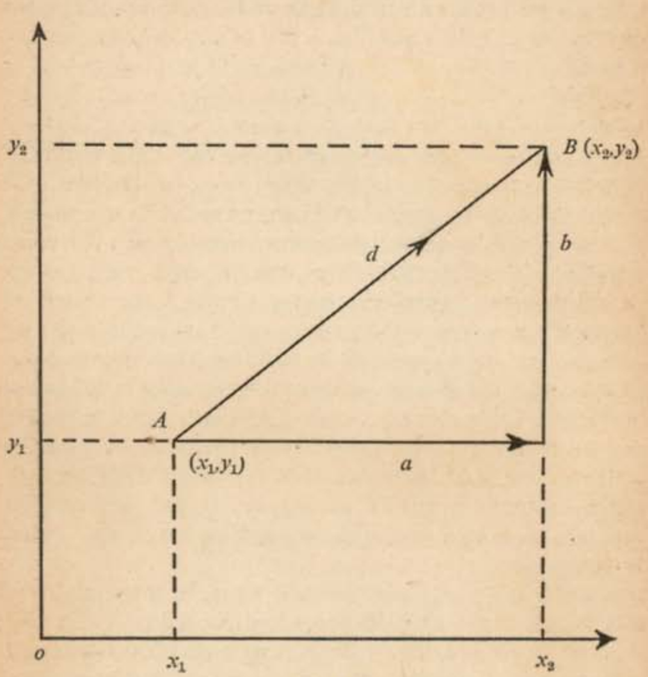


Fig. 1.

The most common frame of reference is, of course, the map grid in an atlas which allows us immediately to give the position of some geographical feature by reading off its latitude and longitude (the co-ordinates).

Obviously, however, the latitude and longitude are not sufficient to uniquely specify a point. For instance, suppose we wish to specify the top of the Empire State Building; in addition to giving the latitude and longitude of the building we should also have to give its height above street level. In other words, we require three co-ordinates to specify a point uniquely; we say, then, that we live in a three-dimensional world.

If, instead of specifying a point in space, we wish to specify an event, then in addition to giving the three spatial co-ordinates we are also required to say when the event took place, i.e. to specify an event we must give four co-ordinates. Thus, at least formally, we can say that we live in a four-dimensional world. However, it is clear that this time co-ordinate is in some very fundamental way different from the three spatial co-ordinates. For while we may, at least in principle, travel to any point in space, we may not travel to any point in time. Starting at any event which has the time co-ordinate 'now', we may only proceed to those events which have time co-ordinates in the future; all past events are inaccessible to us.

While it is clear that if we move from one co-ordinate system to another the spatial co-ordinates of an event will change, it is an assumption of classical physics that the time co-ordinate will remain the same. We shall see in Chapter 3 that in the theory of relativity these distinctions between space and time become somewhat blurred.

To measure the speed with which an object is moving, we measure the distance between two points through which the object passes and the interval of time which elapsed between the object leaving the first point and arriving at the second. Then the average speed of the object as it travels between the two points is defined as the ratio of the spatial interval and the temporal interval. We can then define the speed of an object at a point as the limit of its average speed between two points, as we make the spatial interval between the points smaller and smaller.

If we specify, in addition to the speed of an object, the direction in which it is travelling, we speak of its velocity, i.e. speed is a scalar quantity and velocity is a vector.

In exactly the same way we can define the average rate of change of the velocity of an object, as the ratio of the difference of the velocities at two points, to the time it took to go from one point to another; this is called the average acceleration of the object. The limit of the average acceleration as we bring the two points closer together is called the acceleration of the object at the point. Obviously acceleration is a vector quantity.*

So far we have only discussed the kinetics of our world. To turn this into mechanics we must introduce matter into our considerations. To do this, we require first a measure of the amount of matter in an object; this measure we call the mass of the object. Classically the definition of mass is given in terms of the 'inertia' of the object, i.e. its resistance to a change in its motion. To be more precise, if we define the momentum of an object to be the product of its mass and velocity, then to change the momentum (a vector) we must exert a force on the object. We get our definitions right by equating the force with the rate of change of momentum it produces.† Of course, if the mass of the object is constant, i.e. independent of its velocity as we might intuitively imagine, and as is always assumed in classical physics, then the force is simply equal to the product of the mass and the acceleration.‡

When a force acts on an object and moves it, we say that the force has done work; the work done being equal to the scalar product§ of the two vectors, the force and the distance it moves. The capacity of any system to do work we call the energy of the system. To be exact, what we have described is mechanical energy. As we shall see, there are many different forms of energy and these can be changed from one form into another. But, however we transform the energy, we always find one astonishing fact, the total amount of energy is conserved. There is another conservation law which has been observed, and that is that the total momentum in any closed system also remains constant.

* In Appendix A we present a brief review of the rules for manipulating vectors and the definitions of various terms used in vector analysis. Throughout this book vector quantities will be represented by bold-face letters.

† Newton's Second Law of Motion.

‡ As we shall see in Chapter 3, this velocity independence of mass is only approximately true for velocities much less than the velocity of light.

§ Appendix A.

These conservation laws have a much wider applicability than simply classical physics, and they hold even in modern physics. We shall return in the final chapter to examine their origin and to discuss various other conservation laws.

Let us examine the concept of energy in slightly more detail. Suppose we have a projectile which we fire at a target. When the projectile strikes the target, the target experiences a force which causes it to recoil, i.e. work has been done. The capacity of the projectile to do this kind of work is called its kinetic energy and depends entirely for its existence on the fact that the projectile was in motion.* Suppose, however, we fire the projectile vertically upwards. It will rise to a certain height and then fall back to the surface of the Earth again. If we look at the projectile at the highest point of its journey we find that it is momentarily at rest, i.e. it has lost all its kinetic energy. But if the total energy is conserved, where has it all gone? The answer lies in our neglect of the force which has caused the projectile to come to rest, the force of gravity. The projectile has been working against this force since it left the ground. We say that the kinetic energy has been converted into gravitational potential energy.† This potential energy can, of course, be readily converted back into kinetic energy by simply letting the particle fall back to the ground.

Suppose we were to carry out such an experiment, and we measured the initial and final velocities of the projectile and the height to which it rose, we could check the law of conservation of energy and we should find that it was violated,

* Classically the kinetic energy of an object is equal to half the mass of the object times the square of its velocity. This follows immediately from the definition,

$$\text{K.E.} = \int_0^x \mathbf{F} \cdot d\mathbf{x} = \int_0^x m\frac{d\mathbf{v}}{dt} \cdot d\mathbf{x} = \int_0^{\mathbf{V}} m\mathbf{v} \cdot d\mathbf{v} = \tfrac{1}{2}m\mathbf{V}^2$$

$\mathbf{F}$ = force, m = mass, $\mathbf{v}$ = instantaneous velocity, $\mathbf{V}$ = final velocity.

† If we assume that the force of gravity is constant over the region of motion of the projectile, then the gain in potential energy is simply

$$\text{P.E.} = \mathbf{F}h = mgh$$

$\mathbf{F}$ = force of gravity, m = mass, h = height reached

g = acceleration due to gravity.

i.e. we should find that there is an apparent loss of energy during the experiment. Where, then, has this lost energy gone? If we were very observant we should notice that on its return the projectile was warmer than when it left, as also would be the air through which it has moved. Therefore, heat has been generated by the work done against the air resistance. If we measured this heat energy we should find that it exactly balanced our loss of mechanical energy and restores our conservation law.

Before we move on to discuss further the relationship of thermodynamics to classical mechanics, let us pause and consider this remarkable gravitational force. Newton's universal law of gravity states that every mass in the universe attracts every other mass in the universe with a force which is proportional to the product of the two masses and inversely proportional to the square of the distance between them. Strictly speaking, this law defines a new form of mass, the gravitational mass, as opposed to the inertial mass discussed above. However, the inertial and gravitational masses appear always to be identical, and we shall not draw any distinction between them.

The distinction between weight and mass should now be clear. The weight of a body is a measure of the attraction of that body to neighbouring gravitating masses, e.g. the Earth. If the object is taken far enough away from other objects, it will cease to have any weight (just as an astronaut in orbit becomes 'weightless'). However, it will not have changed its mass, which is simply a measure of the amount of matter in the object.

We can write the law of gravitation as a particularly simple mathematical formula.

$$F = G\,\frac{m_1 m_2}{d^2} \qquad (1.1)$$

where F is the gravitational force, m_1 and m_2 are the gravitating masses, d is the distance between them and G is a constant called the universal gravitational constant equal to $6{\cdot}6 \times 10^{-8}$ cgs units.

In our initial discussion of forces we had a very clear physical picture of the force being transmitted to the target by the projectile, but obviously with gravitational forces we are faced with something entirely new; action at a distance. Two massive bodies may exert a gravitational force on one another with no apparent physical connection between them; what then transmits the force? The short answer is that we do not really know. In the final chapter, where we briefly discuss the general theory of relativity, we shall return to examine this question from a different point of view.

A region of space in which a massive particle experiences a gravitational force we call a gravitational field. The source of the field is, of course, other massive bodies. Can we be more explicit about what we mean by the field? Suppose we have a test particle of mass m which we put at a point in space. If it is in a gravitational field, it will experience a force $\mathbf{F}$ which will produce in it an acceleration $\mathbf{a}$. This acceleration vector is not a function of any property of the test particle, it depends on the position in space at which the particle is placed. If we knew the value of $\mathbf{a}$ at every point in space, we would know exactly the magnitude and direction of the force acting on any particle introduced into the field; it is simply

$$\mathbf{F} = m\,\mathbf{a} \qquad (1.2)$$

where m is the mass of the particle. We talk, then, of $\mathbf{a}(x, y, z)$* being a vector field and call the value of the vector at any point the field strength at that point.

It is sometimes more convenient, instead of ascribing a vector to all points in space, to represent the field by a scalar. Thus, instead of needing three components of a vector at each point in space, we only need a single number. To obtain such a scalar, let us consider the potential energy $U(x, y, z)$ of our test particle in the gravitational field, it is

$$U(x, y, z) = \int \mathbf{F} \,.\, d\mathbf{x} = m \int \mathbf{a} \,.\, d\mathbf{x} = m\, V(x, y, z). \qquad (1.3)$$

---

* By writing $\mathbf{a}(x, y, z)$ we mean that the vector $\mathbf{a}$ is a function of the three spatial co-ordinates, x, y and z.

Once again we see that the scalar quantity $V(x, y, z)$ is not a function of any property of our test particle but depends only on the point in space; V is called the potential of the field. It is clear that to obtain the field strength we need only calculate the gradient* of the potential.

There are, of course, many other potentials besides those due to gravitational fields. Just as a particle can have an intrinsic property which we call its mass, it can also have an intrinsic property which we call its electric charge. However, unlike masses which always attract one another, there are two types of electric charge, called positive and negative. The law of force for charges is of exactly the same form as that for gravitational forces, i.e. the force is proportional to the product of the charges and inversely proportional to the square of the distance between them, but the force is a repulsion in the case of like charges and an attraction in the case of unlike charges.

A charged particle can possess electrostatic potential energy and can be said to move in an electrostatic field in strict analogy with the gravitational case. In this case the force on a charge e is

$$\mathbf{F} = e\,\mathbf{E} \tag{1.4}$$

where E is the electrostatic field strength. Similarly, one could go on and introduce magnetic fields and nuclear fields, etc., and these we shall study in later chapters.

So far we have restricted our discussion to static fields, that is, the field strengths and potentials depend only on the spatial co-ordinates and not on the time. This implies that the sources of the fields are in general stationary. In the next chapter we shall investigate non-stationary fields.

Armed with the laws of classical mechanics and a knowledge of gravitational and electrical forces, we are in a position, at least in principle, to tackle any problem in classical physics. We shall illustrate this by devoting the remainder of this chapter to showing how we can obtain an explanation for a few of the physical properties of matter.

---

* Appendix A.

THE ATOMICITY OF MATTER

By the end of the eighteenth century the existence of a number of distinct chemical elements and compounds had been recognised. An element is defined as a substance which cannot by any further chemical treatment be split into simpler parts; a compound requires the presence of a number of elements, e.g. the gases oxygen and hydrogen are elements, but water is a compound formed by the combination of oxygen and hydrogen.

At first, the elements appeared to have random chemical properties, but slowly order grew out of chaos, and by 1872 Mendeleeff was able to order the elements in a table according to their chemical characteristics. There were many gaps in the original Periodic Table of the Elements, and to fill these gaps Mendeleeff predicted the existence of many new elements together with their chemical properties, all of which were subsequently discovered experimentally.

The first observed fact about any chemical reaction is that the total weight of the reaction agents remains unchanged. In addition to this, there are two very startling regularities in the way in which chemical substances react. Firstly, substances combine only in fixed simple proportions by weight, e.g. one gramme of hydrogen combines with eight grammes of oxygen to form nine grammes of water. Secondly, if two elements can form more than one compound, then the ratios of the weights of the elements in the compounds are simple integer multiples of each other, e.g. the ratio of the mass of oxygen to the mass of hydrogen in hydrogen peroxide is twice the same ratio in water.

These form what are known as the laws of constant and multiple proportions.

In addition, for gases, not only is there a simple rule for the reacting weights but also one for the volumes, e.g. two volumes of hydrogen combine with one volume of oxygen to form two volumes of water vapour, at constant pressure; known as the law of constant volumes.

If elements could exist in arbitrarily sized units, we should not expect such regularities to exist. This led Dalton to suggest that all elements are constructed from basic units, or atoms; small, indivisible, indestructible particles. The atoms of a given element are all identical, but different from the atoms of every other element. Compounds are considered to be built out of molecules; that is, a fixed combination of atoms of different elements. It should be noted in passing that under certain physical conditions of temperature and pressure some elements have a molecular rather than an atomic structure, but in these cases all the atoms in the molecule are identical.

The laws of constant and multiple proportions are immediately explained by assuming that the atoms of an element have a characteristic weight, the atomic weight. To explain the law of constant volumes requires one further assumption, known as Avogadro's hypothesis, which states that equal volumes of gas under the same physical conditions contain equal numbers of molecules.

Along with the idea of an atomic weight we introduce a corresponding molecular weight, being the sum of the atomic weights of the atoms forming the molecule. These weights are not usually measured in pounds or grammes but in units of a standard atomic weight. The standard atom is taken to be carbon, and it is said to have a weight of 12 units. On this scale the atomic weight of hydrogen, the lightest of the elements, is approximately 1, oxygen is 16, while the molecular weight of water is 18. We can explain the oxygen + hydrogen = water reaction by saying that both oxygen and hydrogen are formed out of diatomic molecules (to satisfy the law of constant volumes), and one atom of oxygen combines with two atoms of hydrogen to form one molecule of water.

We also talk about a mole of a substance; this being the amount of the substance contained in its molecular weight in grammes. Thus a mole of oxygen weighs 32 g. It is clear that the number of molecules in a mole of a substance is a constant for all chemical substances and that by the law of constant volumes a mole of a gas will occupy a fixed volume, no matter what the gas. The volume occupied by a mole is found to be

approximately 22·4 litres at normal atmospheric pressure and at a temperature of 0° C. The number of molecules per mole is called Avogadro's number and is usually denoted by N.

If one cools a mole of gas it at first simply contracts, eventually becomes a liquid and finally a crystalline solid (e.g. steam, water, ice). The volume of the solid can be measured and should be approximately N times the volume of a molecule. The distance between molecules can be determined by X-ray diffraction techniques (see p. 82) from which the size of the molecules can be obtained and hence N calculated; the result is that N is approximately $6{\cdot}025 \times 10^{23}$. Since a mole of hydrogen weighs two grammes but contains N molecules, each atom of hydrogen weighs approximately 10^{-24} g. Also, since one mole of a gas usually contracts to a solid volume of the order of its molecular weight in cubic centimetres (e.g. 18 g ice occupy approximately 18 cm^3), we see that the maximum size of a simple molecule is of the order of 10^{-24} cm^3, i.e. it would take 100 million atoms laid side by side to cover one centimetre.

Subsequently it has been discovered that if the elements are ordered according to their atomic weights, then this classification coincides with that of the Periodic Table.

That the atoms have associated with them electric charges was discovered accidentally by Nicholson and Carlisle. Instead of completing an electric circuit by joining two wires forming part of the circuit, the two ends of wire were placed in water, and a gas immediately began to be formed at the negative electrode. This gas was quickly recognised as hydrogen. Faraday examined this new phenomenon and soon discovered that many substances decomposed on the passage of an electric current through them. He recognised that there was a definite quantity of electricity associated with each atom, or ion, taking part in the electrolysis, as it was called. From this, he was able to define a basic unit of electric charge, the coulomb, being that quantity of electricity which liberates 0·001 1180 g of silver in electrolysis. From this we obtain our unit of electric current, the ampere, being the flow of one coulomb of electricity per second.

The study of the motion of the molecules in matter from the point of view of classical mechanics forms the subject material of the kinetic theory of matter. It is clear that we could never hope to follow the motion of every single molecule ($6{\cdot}025 \times 10^{23}$ molecules per mole), and hence we shall have to make approximations and use statistics to analyse the behaviour of aggregates of molecules. We close this chapter by examining a few simple systems and seeing what the kinetic theory predicts for them.

THE KINETIC THEORY OF MATTER

The molecules of a substance are in continual motion and thus possess kinetic energy. It is this energy of the moving molecules which is the form of energy that we know as heat. We can then summarise the difference between solids, liquids and gases as follows.

In a solid, the molecules are localised about certain positions by their strong mutual attraction (electrical in origin). As the temperature rises, the molecules vibrate more and more strongly about these positions until at a critical temperature the vibrations become so strong that they break the bonds which maintain the rigid structure, i.e. the solid has melted, and the substance becomes a liquid. In a liquid, the molecules are free to move at random, but the attractive forces between them still limits them to move within a body of fixed volume. If the temperature is raised still further, the motion of the molecules will become so violent that those near the surface of the liquid will escape. The liquid is now said to be boiling. Above this temperature the substance will expand to fill any volume available to it; it has become a gas.

The simplest system we could imagine would be a gas of non-interacting molecules. The assumption that the interaction between the molecules is negligible is probably reasonable in a low density system where the average spacing between the molecules is large. The molecules will then travel in straight lines at constant velocity until they strike the walls of the vessel containing the gas. If we assume that no heat is conducted away by the walls of the vessel, i.e. that the total energy of the gas is constant, then the collision of a molecule with the wall will have to be elastic. That is, in order to conserve its kinetic energy it will have to rebound into the vessel with the same speed with which it struck the wall.

Let us picture an area of the vessel wall containing such a gas (Fig. 2). Molecules will be continually striking it and rebounding back into the body of the gas. If a molecule of mass m has a component u_1 of its velocity in the direction of the reference axis marked 1 in Fig. 2, which is perpendicular to the wall of the vessel, it will bounce back into the vessel having undergone a change of momentum $2m\,u_1$.

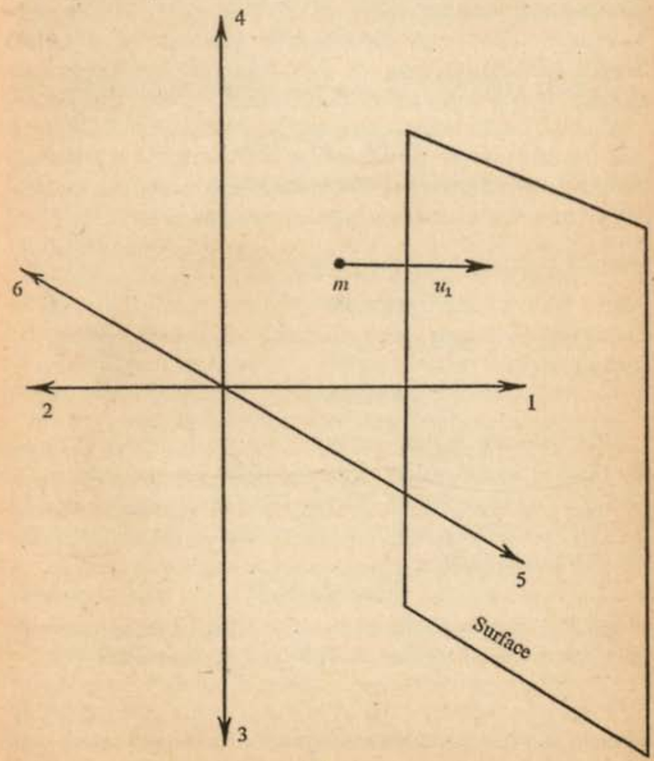


Fig. 2.

Now suppose all the molecules have the same component of velocity in the 1 direction, then the total rate of change of momentum per unit area of the wall of the vessel will be $nu_1 \times 2mu_1 = 2nmu_1^2$ where n is the number of molecules per unit volume. This, then, is the force per unit area on the wall of the vessel, or simply the pressure of the gas. However, there are some assumptions which are not quite realistic. Firstly, not all molecules will have the same component of velocity, and so we replace u_1^2 by its average value denoted by $\langle u_1^2\rangle$. Secondly, only half the molecules will be travelling towards the wall and so we drop the factor 2, leaving us with the expression for the gas pressure

$$P = nm\langle u_1^2\rangle. \tag{1.5}$$

Now all directions of motion are equally probable, and hence

$$\langle u_1^2\rangle = \langle u_2^2\rangle = \langle u_3^2\rangle, \tag{1.6}$$

while the average total velocity squared is

$$\langle u^2\rangle = \langle u_1^2\rangle + \langle u_2^2\rangle + \langle u_3^2\rangle = 3\langle u_1^2\rangle,$$

thus the expression for the pressure becomes

$$P = \tfrac{1}{3}nm\langle u^2\rangle. \tag{1.7}$$

Consider one mole of a gas so that $n = N/V$ where V is the volume, thus

$$PV = \tfrac{1}{3}Nm\langle u^2\rangle. \tag{1.8}$$

Now the total thermal energy of the gas is simply equal to the number of molecules in the gas times the average kinetic energy for each molecule. But the thermal energy of a system is proportional to a quantity which we call its temperature; hence we may write

$$N\tfrac{1}{2}m\langle u^2\rangle = CT \tag{1.9}$$

where C is simply a constant known as the heat capacity of the system. We now combine (1.8) and (1.9) to write

$$PV = NkT \tag{1.10}$$

where k is a new constant called the Boltzmann constant and Nk has the approximate value 2 cal deg^{-1} mole^{-1}.

We see in equation (1.10) a complete statement of the classical gas laws. Firstly, Boyle's Law (1660) which states that for a fixed mass of gas at constant temperature the product of the pressure and the volume are constant. Secondly, Charles's Law (1787) which states that for a fixed mass of gas at constant pressure the volume is proportional to the temperature.

Suppose, now, that the walls of the vessel are moving inwards, then the molecules would gain kinetic energy on collision with the walls, in the same way that a tennis ball gains velocity when hit by a racket. Thus the average kinetic energy of the molecules will rise; but since the average kinetic energy is proportional to the temperature, this means that the temperature will rise. Thus the gas will heat up if it is compressed; this is commonly observed when using a bicycle pump. Correspondingly, if the walls of the vessel move outwards there will be a fall in the average kinetic energy, and hence the temperature will fall and the gas will cool on expansion; this is commonly observed in an aerosol can which cools considerably on use.

Does this mean that there is a net gain of energy when the wall moves inwards and a loss of energy when the wall moves outwards? The answer is no. When we take into account the forces required to move the walls, we find that the total energy is once more conserved. This emphasises once again the need to consider the *whole* system when discussing the conservation laws.

In Appendix B we discuss the statistics of large numbers of particles, and here it will suffice to quote the result for the classical statistics of Boltzmann. In any assembly of particles the number of particles with energy in the range E to $E+dE$ is proportional to $e^{-E/kT}\,dE$. This immediately implies (Appendix B) that each term which enters the energy quadratically will on average contribute $\frac{1}{2}kT$ to the energy of the system, e.g. in the case of the non-interacting gas discussed above the energy of a molecule is simply $\frac{1}{2}m(u_1^2+u_2^2+u_3^2)$ and thus the average energy of the molecule will be $\frac{3}{2}kT$. Hence the specific heat, i.e the heat capacity per mole of such a gas will be simply $\frac{3}{2}Nk \simeq 3$ cal deg^{-1} mole^{-1}. For many monatomic gases at low densities this is, in fact, found to be a good approximation. Suppose that we have a polyatomic gas, then, in addition to the kinetic energy of the molecules, we can also have rotations of the molecules or vibrations of the atoms forming the molecule. These terms also contribute quadratically to the energy, and hence each additional degree of freedom adds $\frac{1}{2}Nk$ to the specific heat. Many gases are found to have specific heats which can be interpreted in this way. However, the agreement only holds over a limited range of temperature and pressure.

Let us leave gases now and consider what might happen in a solid. Here the atoms are localised about certain points, and their only freedom is that they can vibrate about these points. Thus there will be three terms in the kinetic energy and three terms in the potential energy entering the expression for the total energy, and we should expect the specific heat of a solid to be $3Nk \simeq 6$ cal deg^{-1} mole^{-1}. Dulong and Petit over 100 years ago had established that many solids have a specific heat of approximately 6 cal deg^{-1} mole^{-1}, but once again this law breaks down very badly at low temperatures.

Finally, we consider the conduction of electricity in a metal. As a model for a metal we shall consider a set of heavy masses (the atoms) held in fixed positions by their mutual interaction, and around these atoms we shall assume that there is a gas of charged particles, each of charge e and mass m. Again we shall consider the gas to be of such a low density that collisions between the particles are much less important than collisions between the particles and the atoms. If we now apply a constant electric field $\mathbf{E}$ to the metal, the charged particles will all experience a force $e\mathbf{E}$ and will thus be accelerated. Thus a current will start to flow; the density of the current $\mathbf{I}$ being simply $ne\langle \mathbf{v}\rangle$ where n is the number of particles per unit volume and $\langle \mathbf{v}\rangle$ is the average velocity. Now, if there is no decelerating mechanism the average velocity will just increase indefinitely and hence the current will increase indefinitely; this is certainly not observed. However, the charged particles

will collide with the atoms, and if we assume that immediately after each collision the velocity of the particle is completely random, i.e. the average velocity immediately after a collision is zero, then the average deceleration due to the collisions is simply $\langle \mathbf{v}\rangle/\tau$ where τ is the average time between collisions. For a steady current this deceleration must be equal to the acceleration produced by the field $e\mathbf{E}/m$ and thus

$$\mathbf{E} = \frac{m\langle \mathbf{v}\rangle}{e\tau} = \left(\frac{m}{ne^2\tau}\right)\mathbf{I}.$$

This is the well-known Ohm's Law of elementary electricity which states that the current produced in a conductor is proportional to the applied voltage (field strength).

There are many more examples of kinetic theory analysis of the behaviour of matter. The above examples have been chosen at random to indicate how an explanation of the macroscopic behaviour of matter may be obtained from an analysis of the microscopic particles from which it is formed.

However, in all cases the results of a kinetic theory analysis are valid only for a limited range of physical conditions, and always at low temperatures the agreement between theory and experiment is destroyed.

2. Electromagnetism, Light and Waves

THE ELECTROMAGNETIC FIELD

In the first chapter we briefly introduced the notion of electric charges and fields. Let us examine these ideas in some more detail. We know the field due to a static electric charge, it is

$$\phi = -\frac{e}{r} \qquad (2.1)$$

here ϕ is the potential at any point a distance r from the charge e. The electric field strength is obtained by taking the gradient of this potential (Appendix A). Suppose now that I have a number of charges $e_1, e_2, \ldots$, what then is the field? Here nature is very kind to us and presents us with a total field which is simply the sum of all the individual fields; this is known as the Principle of Superposition. For an extended charge distribution of density ρ I can then write

$$\phi(x, y, z) = \int \frac{\rho(x', y', z')}{|\mathbf{r}-\mathbf{r}'|} dx'dy'dz'. \qquad (2.2)$$

We can obtain a very simple pictorial representation of the electric field due to a single charge by drawing 'field lines' emanating radially from the charge, the number of lines being proportional to the charge (Fig. 3). We put arrows on these lines to indicate the direction of the vector field, i.e. the arrow point outwards for a positive charge and inwards for a negative charge. Now the area of a sphere of radius r centred on the charge will be proportional to r^2 and thus, since the total number of lines is constant, the number of lines cutting unit area of such a sphere will be inversely proportional to r^2. We notice an interesting coincidence: the field strength produced by a point charge is radial and inversely proportional to the square of the distance from the charge, our field lines are also radial and their density falls off as the square of the distance. Hence we can use the field lines to give us a pictorial representation of the field in which the arrows indicate the direction of the field and their density indicates the field strength. Clearly we could go on to construct analogous pictures for more complex charge configurations.

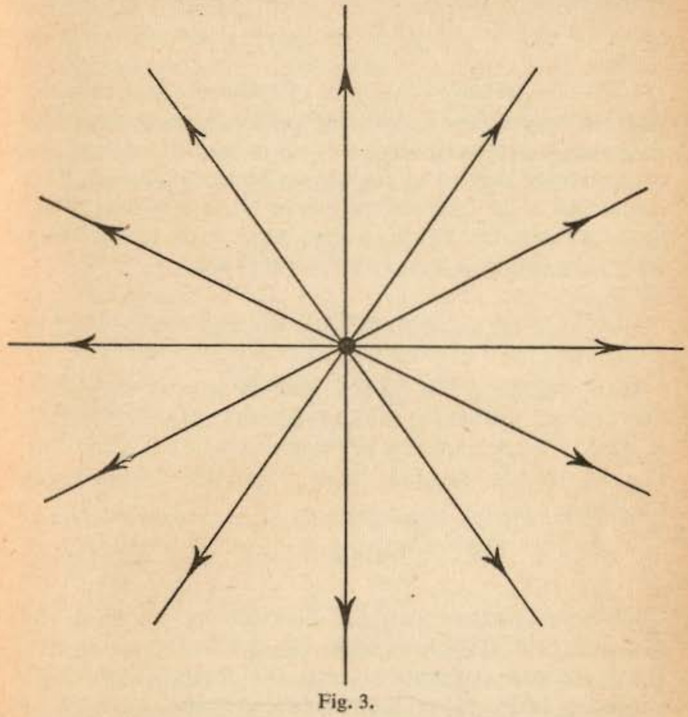


Fig. 3.

In Appendix A we introduce the idea of the divergence of a vector and define it mathematically. But, as its name would suggest, it is simply a measure of the number of lines diverging from a small volume at the point of interest. Clearly, if no charge lies within this small volume the divergence will be zero, since just as many lines will enter as leave the volume. We can choose to draw the number of lines per unit charge so that the divergence of the field is equal to the charge density at that point, or simply

$$\operatorname{div} \mathbf{E}(x, y, z) = \rho(x, y, z). \qquad (2.3)$$

Until now we have considered everything to be static, but suppose we consider a system in which the charges are moving. A charge in motion is what we call an electric current; the current being defined as the charge times its velocity. This means that if the total electric charge is conserved the divergence of the current j from a point is equal to the rate at which the charge density ρ is decreasing at that point, i.e.

$$\operatorname{div} \mathbf{j}(x, y, z) = -\frac{\delta\rho}{\delta t}(x, y, z, t). \qquad (2.4)$$

If the charges are in motion, then in general the electric field strength will vary in time. Equations (2.3) and (2.4) show that the time rate of change of the electric field will act exactly like an electric current; called Maxwell's Displacement Current $\mathbf{j}_d$, i.e.

$$\frac{\delta \mathbf{E}}{\delta t}(x, y, z, t) = \mathbf{j}_d. \qquad (2.5)$$

We now introduce into our discussion a new field, the magnetic field. Gilbert, in 1600, was the first to realise that the Earth was a gigantic magnet and that the directional properties of lodestone (the primitive compass) were due to the Earth's magnetic field. Just as there are two types of electric charge, there appear to be two types of magnetic charges, or poles, as they are called. Because of their directional properties they are called simply north and south poles. But whereas it is possible to produce an isolated negative or positive electric charge, the magnetic poles always go around in pairs; whenever there is a north pole there is a corresponding south pole of equal strength.

It was early realised that there was some very intimate connection between electricity and magnetism. In 1820 Oersted showed that a magnet will always set itself at right angles to a wire carrying an electric current. Shortly afterwards Ampère observed that a small current loop behaves exactly like a magnet. Today it is accepted that the sources of all magnetism are electric currents. We can then see immediately why there is always a north and south pole together, because every current loop has two sides; which is the north side and which is the south depends only on which way the current goes round the loop. The equation analogous to (2.3) which relates the magnetic field strength **B** to a steady current density j is

$$\operatorname{curl} \mathbf{B} = \frac{1}{c}\mathbf{j}\text{*} \qquad (2.6)$$

where c is a constant; its value is chosen so that the electric and magnetic units of charge are the same. If we have a time dependent electric field present, then (2.6) is modified by (2.5) to read

$$\operatorname{curl} \mathbf{B} = \frac{1}{c}\left(\mathbf{j} + \frac{\delta \mathbf{E}}{\delta t}\right). \qquad (2.7)$$

Since north and south poles always occur together and are of equal strength, the net divergence of the magnetic field must be zero, i.e.

$$\operatorname{div} \mathbf{B} = 0. \qquad (2.8)$$

To complete the results necessary for a full discussion of electromagnetism, Faraday, in 1831, showed that if a magnet was moved in the vicinity of a conductor then a current will flow in the conductor. The exact relation between the electric field and the magnetic field is given by

$$\operatorname{curl} \mathbf{E} = -\frac{1}{c}\frac{\delta \mathbf{B}}{\delta t}. \qquad (2.9)$$

Let us collect together all these important equations

$$\operatorname{div} \mathbf{E} = \rho, \quad \operatorname{div} \mathbf{B} = 0 \qquad (2.10)$$

---

[*] In Appendix A we define what we mean by the curl of a vector.

$$\text{curl}\ \mathbf{B}=\frac{1}{c}\left(\mathbf{j}+\frac{\delta\mathbf{E}}{\delta t}\right),\ \text{curl}\ \mathbf{E}=-\frac{1}{c}\frac{\delta\mathbf{B}}{\delta t}.$$

These are known as Maxwell's Equations. They contain all there is to know about electromagnetism and represent perhaps the most beautiful piece of mathematical analysis in the whole of science.

Suppose we have a test charge which we place in an electric field, and we now move the source charges so that the electric field will change. How quickly does it change, i.e. how long is it before the test charge feels the effect of changing the position of the source charges? To answer this, let us take the curl of each side of Equation (2.9). In Appendix A we show that curl (curl $\mathbf{E}$) = grad (div $\mathbf{E}$) $- \nabla^2\mathbf{E}$. But div $\mathbf{E}=\rho$, and if we are interested in the space between the source charges and the test charges then here the charge density is zero. The right-hand side of the equation becomes $-\frac{1}{c}\frac{\delta}{\delta t}$ curl $\mathbf{B}$, but in a space where there are no currents curl $\mathbf{B}=\frac{1}{c}\frac{\delta\mathbf{E}}{\delta t}$. Gathering these results together we have

$$\nabla^2\mathbf{E}-\frac{1}{c^2}\frac{\delta^2\mathbf{E}}{\delta t^2}=0. \qquad (2.11)$$

Equation (2.11) is the standard form of equation describing a wave motion propagating with a velocity c. A similar result holds for the magnetic field $\mathbf{B}$. The constant velocity of propagation of the electromagnetic field c is measured to have the approximate value 3×10^{10} cm s^{-1}.

Maxwell's equations also imply that the electric and magnetic vectors will be perpendicular to each other and to the direction of propagation (Fig. 4).

In 1887 Hertz, stimulated by Maxwell's writings, discovered waves which were definitely of an electrical origin emanating from a region of rapidly changing electric field; these waves we know as radio waves.

Let us make it clear what is happening. Suppose we have an electric charge at rest, it will be surrounded by a static electric field. Now, if the charge is made to oscillate, ripples will propagate in the electric field and also in the magnetic field which is associated with the moving charge, just as when a pebble is dropped in a pool of water ripples radiate out from the centre of the disturbance.

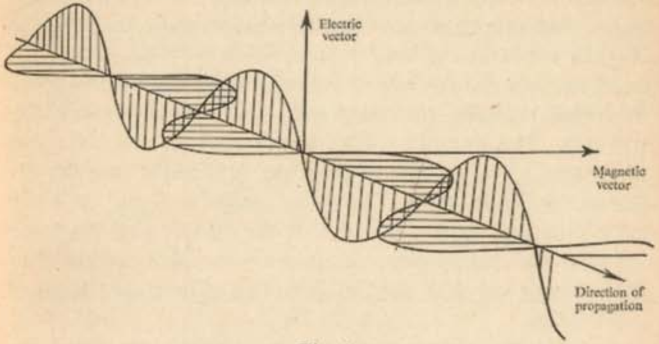


Fig. 4.

Finally, if we have a charge e travelling with a velocity $\mathbf{v}$ it is equivalent to a current $\mathbf{j}=e\mathbf{v}$. Thus the total force on such a particle moving in an electric field $\mathbf{E}$ and a magnetic field $\mathbf{B}$ will be

$$\mathbf{F}=e(\mathbf{E}+\mathbf{v}\times\mathbf{B}) \qquad (2.12)$$

and is known as the Lorentz Force.

WAVES

Having said that the propagation of the electromagnetic field is a wave motion, let us examine some common properties of waves and the vocabulary used to describe them.

Everyday experience presents us with many illustrations of what we describe as waves. Be they water waves or crime waves, there is an underlying concept of some quantity which varies its magnitude in some fairly regular manner with time at a fixed place and which, at any fixed time, has a series of maxima and minima distributed throughout space in some more or less regular pattern.

Basically there are two types of waves. Firstly, we can have a wave pulse in which the source of the wave only exists for a very short period of time; for example, a long rope fixed at one end and held at the other is given a single violent shake, then a disturbance will travel down the rope as a wave pulse. Secondly, we can have a regular wave train which implies a continuous and regular source of the wave; for example, suppose the end of the rope is shaken up and down in a regular fashion, then a train of waves will propagate along the rope. The quantity which is propagated may be a displacement, as in the case of the rope; or it might be a density fluctuation, as in the case of sound waves in a gas; or it can be fluctuations in the strength of the electric and magnetic fields in the case of electromagnetic wave propagation. For the moment we shall content ourselves with a discussion of wave trains.

Analytically, a wave train moving to the right along the x-axis with velocity c can be represented by a function of the form $f(x-ct)$. For, in order to meet our requirements of regularity outlined at the beginning of this section, we must require that at some point $(x-x_1)$ at some time $(t+t_1)$ the function must be unchanged, i.e. $f(x-ct)=f[x+x_1-c(t+t_1)]$. This is certainly true if $x_1-ct_1=0$, or $c=\frac{x_1}{t_1}$. In other words, the disturbance has taken the time t_1 to travel the distance x_1 and hence is moving with the velocity c. Similarly, a wave moving to the left along the x-axis can be represented by a function $f(x+ct)$. We now show that such a function is a solution of the equation (2.11). $\nabla^2 f(x-ct)=\frac{\delta^2 f}{\delta x^2}=\frac{\delta^2 f}{\delta y^2}$ where $y=(x-ct)$ and $\frac{\delta^2 f}{\delta t^2}(x-ct)=c^2\frac{\delta^2 f}{\delta y^2}$.

Thus
$$\nabla^2 f(x-ct)-\frac{1}{c^2}\frac{\delta^2 f}{\delta t^2}(x-ct)=0$$
as required.

The most important type of wave motion is the simple harmonic wave in which the periodic variations follow the

sine or cosine law (see Fig. 5 where we illustrate a snapshot of the wave), i.e. $t=$ constant. The sinusoidal wave is expressed as $f=A \sin 2\pi/\lambda(x-ct)$. The distance between successive maxima in Fig. 5 is called the wavelength and is denoted by the Greek letter λ. The value of f at its maxima is A, and this is called the amplitude of the wave. If we fix our attention at the point A_1 in Fig. 5, then the value of f will vary with time. The time required for the value of f to fall from A to its minima, $-A$, and then rise back to A again is called the period of the wave and is clearly equal to λ/c. Its reciprocal c/λ is called the frequency of the wave and is denoted by ν.

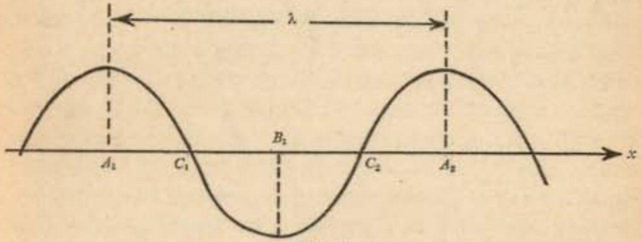


Fig. 5.

If we have more than one wave, then the total displacement is obtained by superimposing the displacements which would result from each wave on its own; this is again simply the superposition principle. For example, we could have the superposition of two waves travelling in opposite directions along the x-axis with the same velocity c. If, as in (2.13), they have the same amplitude then

$$f=A \sin \frac{2\pi}{\lambda}(x-ct)+A \sin \frac{2\pi}{\lambda}(x+ct) \qquad (2.13)$$
$$=2A \sin \frac{2\pi}{\lambda}x \cos 2\pi\nu t.$$

In this case there is no net propagation; we simply have the fixed form illustrated in Fig. 5 but with an amplitude varying periodically in time from $+2A$ to $-2A$. In this case we talk of a stationary or standing wave. The fixed points at which the disturbance is always zero, i.e. points C_1, C_2, etc. in Fig. 5, are called the nodes of the wave. The intermediate points at which the disturbance has a maximum, though fluctuating, value are called the antinodes.

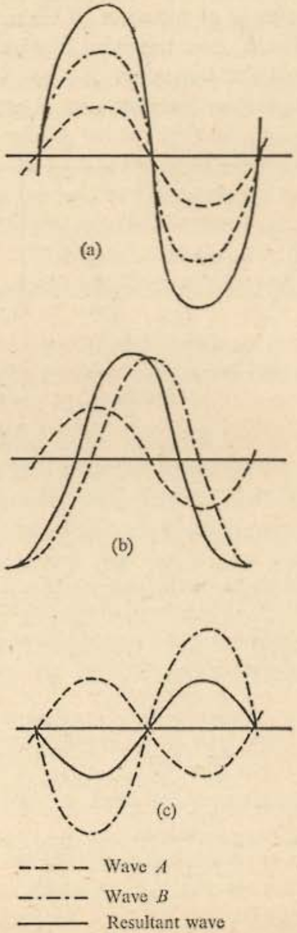


Fig. 6.

Now consider two waves which we shall call *A* and *B* and which, for simplicity, we shall assume have the same wavelength (Fig. 6).

In Fig. 6(*a*) the nodes of *A* and *B* coincide and, moving away from a node, *A* and *B* both either rise or fall below the axis together. In such a case the waves *A* and *B* are said to be in phase. In Fig. 6(*b*) the nodes of the wave *B* have been shifted to the right of the nodes of the wave *A*, and the waves are now said to be out of phase. Finally, in Fig. 6(*c*) the nodes of *B* have again been shifted further to the right until they again coincide with those of *A*. But now, whereas as we move away from a node, *A* will rise above the axis, *B* will fall below the axis; the waves are now said to be in antiphase. In all three cases one obtains a resultant wave (represented by the heavy dark line in Fig. 6) which has for its distance from the axis at any point the algebraic sum of the distances of *A* and *B* from the axis at that point, distances above the axis counting positive and those below the axis negative. This wave *R* is said to be the result of the interference of the waves *A* and *B*. It has its maximum amplitude when the waves are in phase and falls to a minimum when the waves are in antiphase. In fact, two identical waves in antiphase will exactly cancel each other out.

PROPERTIES OF LIGHT

Firstly, we note that light travels in straight lines and that it can be reflected and refracted. An easily proved fact about reflection is that the angle of incidence is equal to the angle of reflection (see Fig. 7). Refraction takes place when light crosses the boundary between media of different density. As the light enters a denser medium it bends so as to be closer to the normal (see Fig. 7). This phenomenon is commonly observed when we look at a stick which is partially immersed in water; the stick appears to have a sharp bend in it where it enters the water. The law describing the refraction of light is known as Snell's Law and states that for any two materials the ratio of the sine of the angle of incidence to the sine of the angle of refraction is a constant. In the case of light entering a material from a vacuum, we call this constant the refractive index for the material. The refractive index n is always greater than 1 and for most glasses has a value of approximately 1·5

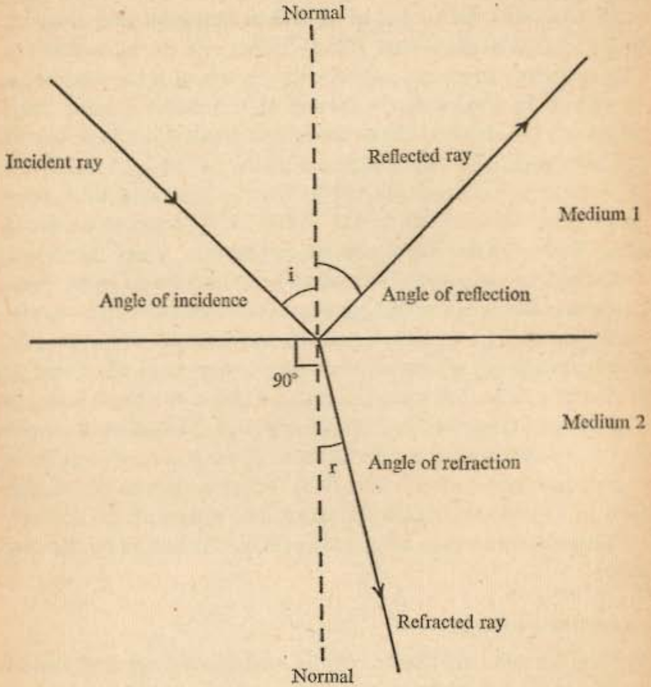


Fig. 7.

$$n = \frac{\sin \hat{\imath}\ (\text{in vacuum})}{\sin \hat{r}\ (\text{in material})}. \tag{2.14}$$

The next most easily observed feature about light is that it exists in different colours. The index of refraction for a given material depends slightly on the colour of the light; we call this dependence dispersion. Some materials are very much more dispersive than others, e.g. air shows very little dispersion, while lead glass exhibits it to a considerable degree.

Newton showed that when a beam of white light fell on a glass prism, behind which was placed a white screen, then a band of colours which he called a spectrum appeared on the screen (see Fig. 8). This is again a commonly observed phenomenon in the form of a rainbow, where the white sunlight, instead of striking a glass prism, is being dispersed by small water droplets in the atmosphere.

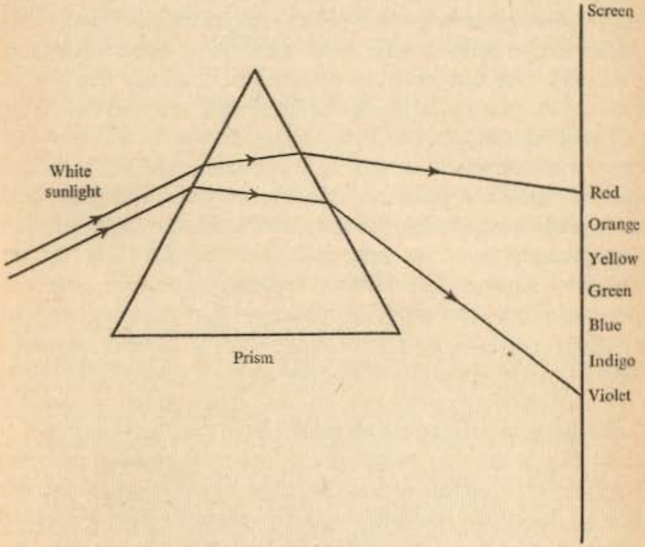


Fig. 8.

If a second prism is now placed in the spectrum, but upside down relative to the first, then the spectrum is recombined and white light emerges.

Should light of a particular colour be allowed to pass through the screen, it is found that it cannot be split up into any further colours.

Finally, we briefly mention the polarisation of light. This can occur either when light is reflected or refracted, or when light is passed through certain crystals. The most convenient material for studying polarisation is called 'polaroid'; this comes in sheets bearing a reference mark in the form of an arrow. An ordinary beam of light will be partially transmitted through a polaroid; however, its ability then to be transmitted through a second polaroid depends on the relative orientation of the reference marks on the two sheets. Maximum transmission occurs when the reference marks are parallel, and falls to zero when they are at right angles, i.e. the light will not be fully transmitted through the second polaroid unless the reference mark of this polaroid lies in the plane containing the direction of propagation of the light and the reference mark of the first polaroid. Light which has passed through a polaroid is said to be plane polarised, and the plane of polarisation is the plane defined by the direction of propagation of the light and the reference mark on the polaroid. Light reflected from a transparent surface, such as a sheet of glass, at an angle of incidence $\hat{\imath}$ such that $\tan \hat{\imath} = n$, the refractive index, is also plane polarised, the plane of polarisation being the plane containing the incident and reflected rays. This reflected ray will not be fully transmitted through a polaroid unless the reference mark lies in the plane of polarisation. Similarly, this reflected ray will not be fully reflected from another surface unless the plane of polarisation for this second reflection coincides with the plane of polarisation of the ray. The ray transmitted through the surface, i.e. the refracted ray, will also be plane polarised, but its plane of polarisation will be perpendicular to that of the reflected ray.

Obviously, sunlight transmitted through the atmosphere will be refracted as it enters from the vacuum of space and will thus be partially plane polarised. Hence we can make sun-glasses out of polaroid with the reference mark horizontal, and these will not transmit a considerable portion of the sunlight.

If a bar magnet is placed parallel to a beam of plane polarised light, then the plane of polarisation will rotate. This magnetic rotation is known as the Faraday Effect, after its discoverer.

THE NATURE OF LIGHT

In order to explain the simply observed properties of light discussed in the previous section, Newton tentatively suggested that a light beam might consist of a stream of tiny corpuscles emitted by a luminous body. Then the fact that light propagates in straight lines follows from the conservation of momentum. The law of reflection can be interpreted as an elastic collision between the corpuscles and the mirror, and follows from the conservation of energy and momentum.

To explain the refraction of light, Newton postulated that the corpuscles were attracted to media of higher density and consequently moved faster in them. If c is the velocity of light in a vacuum and c' the velocity in some given medium, then the index of refraction for that medium is simply the ratio of these velocities

$$n = c'/c. \tag{2.15}$$

It follows at once that n is greater than 1.

To account for the different colours of light, Newton proposed that there be many different kinds of corpuscles. The corpuscles associated with a given colour were all to be identical, but different from the corpuscles associated with any other colour. By assuming that the corpuscles associated with different colours were attracted to denser materials to varying degrees, he was able to explain the observed dispersion of light.

We are then left with the problem of explaining the polarisation of light. This requires that the corpuscles can recognise certain planes as being distinct from others. This cannot happen if the particles are spherical, because then all planes containing the direction of propagation will appear identical. However, should the particles be ellipsoids which are propagated perpendicular to their major axis, we can use the direction of propagation and the orientation of the major axis to define the plane of polarisation. If in a beam of light the orientations of the major axes are random, we talk of unpolarised light; while if the axes are all parallel, the beam is called plane polarised.

For a considerable period Newton's stature as the foremost natural philosopher of his time led to the almost unquestioning acceptance of his ideas relating to the nature of light. What had, for him, been a tentative suggestion became in the hands of his followers a dogmatic statement of fact.

A contemporary of Newton was the Dutch physicist Huygens, who proposed that instead of a stream of particles light consisted of a wave motion. But if a wave motion, then a wave in what? At that time all known wave motions required a medium through which to propagate. However, light can cross a vacuum, e.g. the light from the Sun crosses some 93 million miles of virtual vacuum to reach us. To overcome this problem, the existence of an omnipresent ether was postulated as the medium through which all light is propagated.

It took over a century for Huygen's theory to gain any substantial support, but within a very short period at the beginning of the nineteenth century conversion to the wave theory of light was virtually complete. This conversion was due to a series of observations which seemed to completely refute forever the corpuscular nature of light.

Thomas Young, in 1801, published his *On the theory of light and colours*, in which he described a number of newly observed phenomena in terms of a wave motion for the propagation of light. Firstly, there was the diffraction of light. If a broad beam of light is shone on to an extremely narrow slit in an otherwise opaque screen, the light which passes through the slit, instead of appearing as a narrow beam, spreads out in all directions (see Fig. 9). A similar effect may be noted in water waves. If a set of parallel wave fronts break on an obstacle containing a narrow aperture, then on the far side of the obstacle waves will spread out from the aperture in all directions.

Now, suppose this diffracted beam falls on a second screen containing two parallel slits (Fig. 9) and the diffracted beams from these slits in turn fall on a third screen. Then on this third screen there appears a pattern of light and dark bands for which there is no explanation in the corpuscular theory. The explanation is simple in terms of the interference between

the wave fronts emitted from the slit A and those from the slit B on wave theory of light. Due to the different path lengths

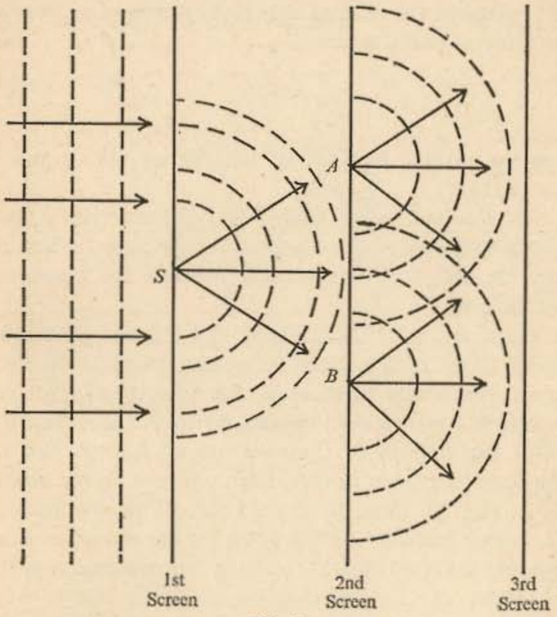


Fig. 9.

from slit A to a point on the third screen and from the slit B to the same point, there will be alternating points on the screen where the waves from A and B will be in phase, hence a light band, and in antiphase, hence darkness.

As to the different colours of light, Young associated these with different wavelengths and from this was able to explain dispersion and the existence of colours in thin films, such as may be seen on a thin layer of oil.

The final blow to the corpuscular theory of light came with a detailed study of the nature of refraction. The wave theory postulated that light experiences a resistance to its motion in denser media, and hence a wave when falling on the interface between two media will be bent in the same way as a row of marching soldiers will turn if those on the inside of the bend march more slowly than those on the outside. The refractive index is given as the ratio of the velocity of light in a vacuum to the velocity in the medium,

$$n = c/c'. \tag{2.16}$$

This is the inverse of what is required by the corpuscular theory (2.15).

In 1850 Foucault performed the crucial experiment, determining that the velocity of light through water is less than that in air and hence apparently refuting the corpuscular theory for ever.

An early value for the velocity of light was obtained by Fizeau in 1849, using a purely terrestrial method. Previously, attempts were made to measure the velocity of light from astronomical observations; notably by Romer in 1675 from a study of the eclipses of the satellites of Jupiter. We shall briefly describe Fizeau's experiment, which is shown schematically in Fig. 10. Light from the source S passes through a lens L_1 which focuses it to the point P after reflection from a transparent glass sheet M_1. The light is made into a parallel beam by the lens L_2, then focused on to the mirror M_2 by lens L_3 which reflects it back through the sheet M_1 to a lens L_4 which focuses the light at the observer O. A toothed wheel is added in such a manner that its perimeter passes through the point P. As the wheel rotates, light is alternately allowed to pass P through a gap between the teeth, or is stopped by a tooth according to the position of the wheel. The experiment consists of rotating the wheel at increasing speed until no light is seen by the observer. This implies that the light which passed through a gap at the point P has been confronted by a tooth on its return journey. If there are N teeth and the wheel is rotating at ω revolutions per second, the time taken for a tooth to move into a neighbouring gap is $1/2N\omega$. This must also be the time which the light has taken to travel from P to M_2 and back to P. If the distance P to M_2 is L, then this

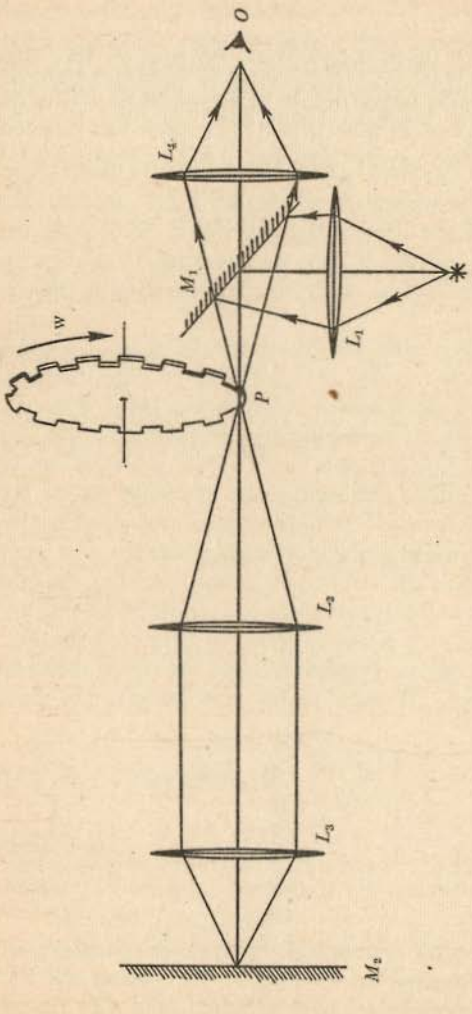


Fig. 10.

time is $2L/c$ where c is the velocity of light, and thus $1/2N\omega = 2L/c$ or

$$c = 4LN\omega. \tag{2.17}$$

Many much more sophisticated versions of this apparatus have been developed since Fizeau's original measurements. The current value of the velocity of light is $2{\cdot}997\,93 \times 10^{10}$ cm s^{-1}.

In 1862 Maxwell pointed out the incredible agreement between the measured velocity of light and the calculated velocity of the electromagnetic waves, and concluded that light was in fact simply an example of electromagnetic radiation but of a much shorter wavelength than Hertz's radio waves.

If matter is built out of electrically charged ions, as suggested by electrolysis, then we have an explanation of why objects glow when heated. As we heat a piece of material, the motion of the molecules increases in energy, i.e. the frequency of the vibrations increases, and since the molecules are charged they will radiate electromagnetic radiation. When the temperature (frequency) is high enough (or, what is the same thing, the wavelength is short enough) for the electromagnetic waves to enter the visible part of the spectrum, the body will glow a dull red. If we continue to heat the body, the frequency of the vibrations increases until the body may become white hot, i.e. all colours (wavelengths) of the visible spectrum have been excited. At even higher temperatures the body will start to give off X-rays, but we postpone discussion of these until a later chapter. In Fig. 11 we present the complete electromagnetic spectrum as given by Maxwell's theory.

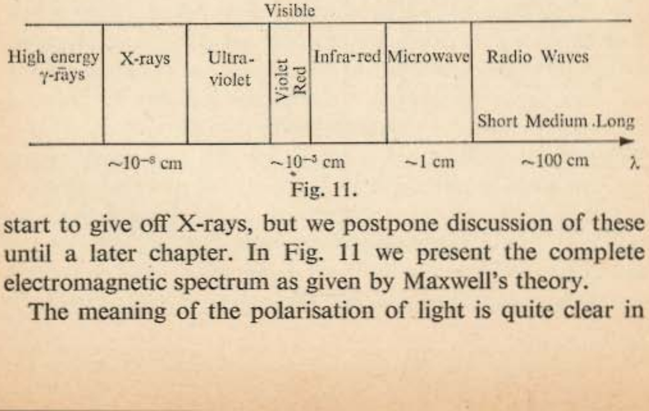


Fig. 11.

The meaning of the polarisation of light is quite clear in the context of the electromagnetic theory. The plane of polarisation is simply the plane of fluctuations of the electric field (see Fig. 4).

Many physicists towards the end of the nineteenth century believed that a combination of the kinetic theory of matter and Maxwell's electromagnetic theory would prove sufficient to explain all the observed phenomena of nature. That it was not to be so, we shall see in the remainder of this book.

3. Relativity

THE MICHELSON-MORLEY EXPERIMENT AND NEWTONIAN RELATIVITY

The ancients believed that the Earth was the centre of the universe and that the Sun, Moon and planets encircled it. Later, Copernicus, Kepler and Galileo had argued that the Sun was the centre of the universe and that the planets followed their courses around it. In time, it was realised that even the Sun held no unique place in the universe and was merely another star wandering through space. But surely there was some fixed centre of the universe, or, what is the same thing, some absolute frame of reference relative to which all the stars in their courses could be measured? What, then, physically was to provide this absolute reference frame? During the nineteenth century physicists came to believe that the ether should provide such a frame. The ether was omnipresent. It was the means by which light travelled from one star to another, and the feeling grew that no single star could be the origin of such an omnipresent material and that if anything could be considered as at absolute rest then it should be the ether.

In 1881 the American scientist Michelson attempted to measure the velocity of the Earth as it moved through the ether. No motion relative to the ether could be detected. However, the accuracy of the experiment was questioned, and in 1887, with the collaboration of Morley, he repeated the experiment with much greater accuracy, but again no motion relative to the ether was discovered. It was this rather startling discovery which was to have such a revolutionary effect on the history of scientific thought.

In essence, the experiment was very simple. Just as the driver of an automobile will know that his vehicle is in motion

by the air resistance that he experiences when he puts his hand out of the window, so a light beam should be able to detect the analogous ether wind as the Earth moves in its orbit. However, before we discuss the experiment in greater detail let us examine the classical ideas on relative motion.

Newton had been rather confusing on the subject of relativity; while saying that translatory motion involved going from one absolute point to another absolute point, he did not define what he meant by an absolute point. Elsewhere he stated clearly that only motion relative to other material bodies can be detected. Presumably he believed that there was some material body which could be said to be at absolute rest, viz. the ether, and thus could be used to define absolute points.

When discussing time Newton was much more explicit. 'Absolute, true and mathematical time, of itself and by its own nature, flows uniformly on, without regard for anything external.' One immediate consequence of this is that all frames of reference must have the same time scale.

Let us consider two frames of reference (S and S' in Fig. 12) which, at some instant of time which we shall call the beginning of time, have their origins coincident. We shall choose the axes of the two systems also to coincide at this time. Now suppose that the system S' moves so that its origin travels at a constant velocity u along the x-axis of the system S.

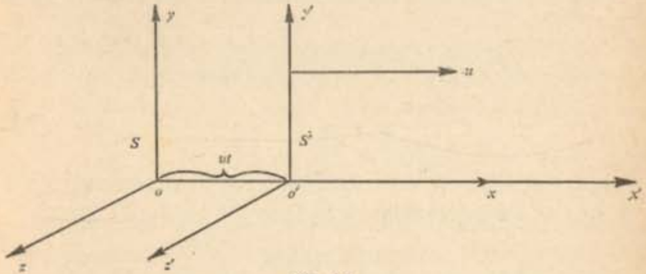


Fig. 12.

We may, for instance, think of O as being a point in a railway station with the x-axis as the railway track and O' being a point on a train moving down the track with constant velocity u.

Now suppose we have an event which in the system S can be labelled by the co-ordinates (x, y, z) and the time t. In the moving system S' will have the co-ordinates (x', y', z') and the time t'. As we have already stated, the time scale is the same in all reference frames and hence

$$t=t'. \tag{3.1}$$

But what is the relationship between the spatial co-ordinates (x, y, z) and (x', y', z')? From Fig. 12 it is obvious that

$$\begin{aligned} x'&=x-ut, \\ y'&=y, \\ \text{and} \qquad z'&=z. \end{aligned} \tag{3.2}$$

The equations (3.1) and (3.2) are said to form a Galilean transformation, and govern the description of the motion of one system relative to another in Newtonian relativity.

Suppose now that there is a point in the system S' which is moving with a constant velocity V' relative to O'. By the definition in Chapter 1 the x-component of the velocity V_x' is

$$\begin{aligned} V_x'&=\frac{x_2'-x_1'}{t_2'-t_1'}=\frac{x_2-x_1-(ut_2-ut_1)}{t_2-t_1} \\ &=\frac{x_2-x_1}{t_2-t_1}-u. \end{aligned} \tag{3.3}$$

But $(x_2-x_1)/(t_2-t_1)$ is by definition the x-component of the velocity of the point relative to O in the system S. Thus

$$\begin{aligned} V_x&=V_x'+u, \\ \text{and similarly} \qquad V_y&=V_y', \\ V_z&=V_z'. \end{aligned} \tag{3.4}$$

The equations (3.4) give the rules for the addition of velocities in Newtonian relativity. Thus a man walking down the corridor of a train will be moving relative to the ground

with a velocity which is the sum of his walking velocity and the velocity of the train. All this is indeed very elementary. But let us see what results it predicts for the Michelson-Morley Experiment.

The experiment consisted of studying the time light took to travel the same distance in two perpendicular directions (Fig. 13(*a*)). Light from a source S is shone on to a semi-silvered mirror M set at an angle of 45° to the incident beam. Some of the light passes through M and travels a distance l to the mirror B, where it is reflected back to M and from M to the observer. The light which does not pass through M is reflected to the mirror A, which is also at a distance l from M, where it is reflected back through M to the observer. Since all the light initially came from the same source S, it was initially all in phase. If the apparatus is at rest, both light beams will have travelled exactly the same distance and taken exactly the same time to do the journey and hence will arrive at the observer still in phase.

Now let us consider the apparatus moving relative to the ether with a constant velocity v, which for simplicity we shall assume is along the direction MB. Then the ray going from M to B will have to battle against the ether wind and will have its velocity reduced to $c-v$, where c is the velocity of light in the ether at rest. On the way back to M from B, however, it will have the wind behind it and hence its velocity will be $c+v$. And so the total time taken for the light to travel from M to B and back again to M, which we shall call t_1, is

$$t_1=\frac{l}{c-v}+\frac{l}{c+v}=\frac{2l/c}{1-v^2/c^2}. \tag{3.5}$$

Meanwhile, the ray going from M to A will have a slightly longer journey to make, because the mirror A will have moved from its initial position A_1 to A_2 by the time the light reaches it from the mirror M (Fig. 13(*b*)). Similarly, the mirrors will move to positions A_3 and M_3 during the return journey of the ray from A to M. The distance M_1 to M_3 is simply vt_2 where t_2 is the time required for the light to make the round trip from M to A and back to M, and thus $M_1M_2=M_2M_3=\frac{1}{2}vt_2$;

the distance M_2A_2 is, of course, fixed at l. If we apply Pythagoras's well-known theorem, about the square of the hypotenuse of a right-angled triangle being equal to the sum of the squares on the other two sides, to the triangle $M_1A_2M_2$, we have

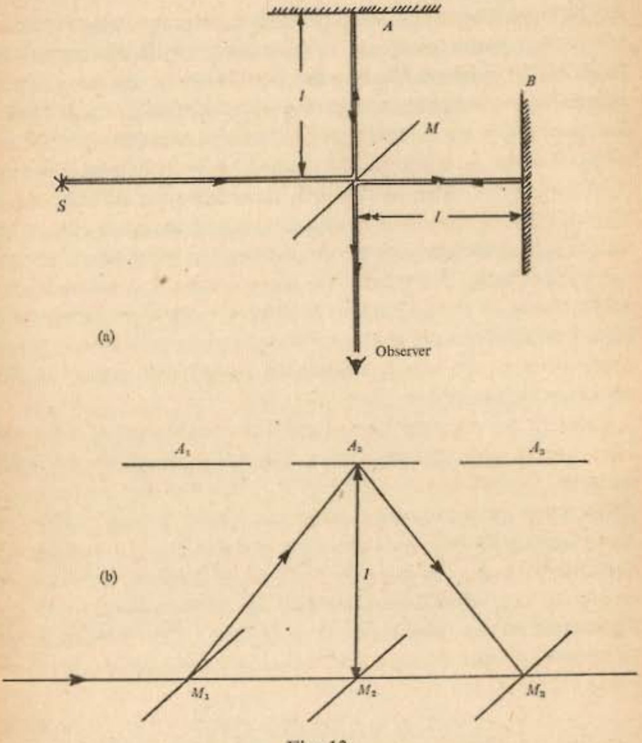


Fig. 13.

$$(M_1A_2)^2=(M_1M_2)^2+(M_2A_2)^2=\tfrac{1}{4}v^2t_2^2+l^2. \qquad (3.6)$$

Thus the total distance travelled by this ray $M_1A_2+A_2M_3$ is $2\{l^2+\frac{1}{4}v^2t_2^2\}^{\frac{1}{2}}$ since the velocity of the ray is c, the time for the trip t_2 is

$$t_2=2/c\{l^2+\tfrac{1}{4}v^2t_2^2\}^{\frac{1}{2}} \qquad (3.7)$$

from which we obtain

$$t_2=2l/c\{1-v^2/c^2\}^{\frac{1}{2}}. \qquad (3.8)$$

Comparing (3.8) and (3.5) we see that

$$t_1=t_2\{1-v^2/c^2\}^{\frac{1}{2}} \qquad (3.9)$$

i.e. the time taken by the second ray is less than the time taken by the first ray. The difference in the time increases as v increases, reaching a maximum when $v=c$ at which stage the mirror B is moving away from the source so quickly that the light never reaches it. If $v=0$ then, of course,

$$t_1=t_2. \qquad (3.10)$$

The fact that t_1 does not equal t_2 in general means that the two rays should reach the observer out of phase. The observer can then use the difference of the phases of the two light beams to calculate how fast his apparatus is moving relative to the ether.

The apparatus was mounted on a massive stone disc, floating in mercury so that it could easily be turned. The total path that each ray had to travel was 22 metres. The experiment was repeated with the apparatus at all possible orientations and at different times of the year, to allow for the different directions of motion of the Earth in its orbit. In all cases $t_1=t_2$, i.e. no motion relative to the ether could be detected. Since the Earth is surely not the centre of the universe, there must be some motion relative to the ether, then what is wrong with (3.9)?

The Irish physicist Fitzgerald and the famous German scientist Lorentz both put forward the same possible explanation independently. The ether was the medium through which electromagnetic waves are propagated, and the force between molecules was known to be electromagnetic in origin; then could it not be that as matter moved through the ether the intermolecular electromagnetic forces underwent some change due to the motion which produced a contraction of the object in the direction of the motion? If this contraction was of the order $\sqrt{(1-v^2/c^2)}$, then the path of the first light ray would be shortened just sufficiently to produce the necessary equality of the two times. Naturally this contraction could never be measured, because any measuring device used would suffer an exactly similar contraction.

While this Fitzgerald-Lorentz contraction answered the apparent paradox posed by the result of the Michelson-Morley Experiment, it was somehow unsatisfactory. It appeared a rather *ad hoc* assumption which had no fundamental justification. The justification was given by Albert Einstein in the form of his Special Theory of Relativity.

EINSTEIN'S SPECIAL THEORY OF RELATIVITY

In 1905 Einstein published his famous paper on the *Special theory of relativity*. 'Special' because it applied only to inertial frames of reference; that is, frames of reference which are not in accelerated motion but move uniformly relative to one another, e.g. S and S' of Fig. 12. Another way of saying the same thing is that the theory only applies to frames of reference in which Newton's laws of motion are valid.

Einstein founded his theory on two postulates. Firstly, the laws of physics are the same for any two inertial frames. This he considered to be obvious from physical experience. For example, if we return to the analogy of the train and the station, this first postulate states that any physical experiment carried out in the station will produce exactly the same result that would have been obtained had the experiment been carried out in the train moving with constant velocity. Secondly, the velocity of light is a constant and independent of the motion of its source, i.e. equations (3.4) do not apply to light signals and hence motion relative to the ether can never be detected. This Einstein took to be established by the Michelson-Morley Experiment.

If we are going to insist on the constancy of the velocity of light and impose it as a restriction on our dynamics, we must give up some restriction which we have hitherto assumed in order to make way for this new condition. The restriction which Einstein relaxed was the idea that there was some absolute time scale. However, he maintained the assumption

that the transformation from the co-ordinates of one system to the co-ordinates of a system moving uniformly with respect to it is linear, i.e. if we make two measurements in one system – for instance, of length – and the first measurement is double the second, then in the new system, although the measurements will have changed, the first length will still be double the second. Thus for the situation shown in Fig. 12 we have

$$\begin{aligned} x' &= \alpha x + \beta t \\ y' &= y \\ z' &= z \\ t' &= \gamma t - \delta x \end{aligned} \qquad (3.11)$$

where α, β, γ and δ are numbers which are to be determined by the condition of the assumed constancy of the velocity of light. However, we notice that the first of these equations can immediately be simplified because the origin of S', i.e. $x'=0$, is always seen in the frame S to be the point $x=ut$, and thus we have

$$\alpha ut + \beta t = 0$$

or

$$\beta = -\alpha u. \qquad (3.12)$$

Thus the first of the equations (3.11) becomes

$$x' = \alpha(x - ut). \qquad (3.13)$$

A simple application of Pythagoras's Theorem shows us that the equation connecting points on the surface of a sphere, having the origin as its centre, of radius r is

$$x^2 + y^2 + z^2 = r^2. \qquad (3.14)$$

Let us now consider the wave front of a light signal emitted by a point source. The wave front will spread out uniformly in all directions about the source, which we shall take to be the origin of the system S in Fig. 12, i.e. it will be a sphere. The radius of the sphere at any time will be the distance that the light signal has gone in that time, thus

$$r = ct. \qquad (3.15)$$

Hence any point on the wave front will be described by the equation

$$x^2+y^2+z^2-c^2t^2=0. \tag{3.16}$$

However, if this signal is to appear the same in all systems, in particular in the frame of reference S', a point on the wave front must also be described by

$$x'^2+y'^2+z'^2-c^2t'^2=0. \tag{3.17}$$

Equating (3.16) and (3.17) and inserting (3.11) and (3.13), we obtain

$$\begin{aligned} x^2+y^2+z^2-c^2t^2&=x'^2+y'^2+z'^2-c^2t'^2\\ &=\alpha^2(x-ut)^2+y^2+z^2-c^2(\gamma t-\delta x)^2 \end{aligned} \tag{3.18}$$

or simply

$$x^2+c^2t^2=(\alpha^2-c^2\delta^2)x^2+2(u\alpha^2-c^2\delta\gamma)xt+(\alpha^2u^2-c^2\gamma^2)t^2. \tag{3.19}$$

By comparing both sides of (3.19) we see that*

$$\begin{aligned} \alpha&=(1-u^2/c^2)^{\frac{1}{2}}\\ \delta&=u\alpha/c\\ \gamma&=\alpha. \end{aligned} \tag{3.20}$$

Thus we finally obtain the equations (3.11) in the form

$$\begin{aligned} x'&=(x-ut)(1-u^2/c^2)^{-\frac{1}{2}}\\ y'&=y\\ z'&=z\\ t'&=(t-ux/c^2)(1-u^2/c^2)^{-\frac{1}{2}}. \end{aligned} \tag{3.21}$$

This is the Lorentz transformation and should be compared with the Galilean transformation (3.2). Firstly, we notice

---

* Comparing the coefficients of x^2 on both sides of (3.19) gives

$$1=(\alpha^2-c^2\delta^2)$$

while the coefficient of xt gives

$$0=2(u\alpha^2-c^2\delta\gamma)$$

and the coefficients of t^2 yield

$$-c^2=(\alpha^2u^2-c^2\gamma^2).$$

These three equations for the three unknown quantities α, δ and γ can then be rearranged to give the result (3.20).

that if the system S' is moving slowly with respect to the system S compared to the velocity of light, i.e. u is very much smaller than c, then u^2/c^2 is very small; for example, if u is 30 miles per hour, then u^2/c^2 is approximately 0·000 000 000 002, thus $\sqrt{(1-u^2/c^2)}$ is to all intents and purposes 1. If we are considering distances of only a few miles, then also ux/c^2 becomes negligible compared to t and we see that (3.21) then reduces to exactly (3.2). However, for systems travelling with high velocities (3.21) predicts behaviour which is considerably different from that predicted by Newtonian physics. Let us consider two of these predictions.

Firstly, suppose there are two points on the imaginary railway track in Fig. 12 which we denote by x_2 and x_1. Then, if in the system S we measure the distance between them and find it to be L, we have the identity

$$L=x_2-x_1.$$

However, if at the same time we measure the distance between the same two points in the system S' we will find a length L' satisfying

$$L'=x_2'-x_1'.$$

Making use of (3.21), we have

$$\begin{aligned} L'=x_2'-x_1'&=\frac{x_2-ut}{\sqrt{(1-u^2/c^2)}}-\frac{x_1-ut}{\sqrt{(1-u^2/c^2)}}\\ &=\frac{x_2-x_1}{\sqrt{(1-u^2/c^2)}}=L/\sqrt{(1-u^2/c^2)}. \end{aligned} \tag{3.22}$$

Thus the length of the track between the two points will appear shorter if we measure them from the train than if we measure them from the platform. Not only that, but the faster the train is travelling the shorter the distance between the two points will appear to someone on the train, until when travelling at the velocity of light the distance between the points will be zero.

Now, of course, this contraction is just the same contraction which Fitzgerald and Lorentz postulated on an *ad hoc* basis to explain the null result of the Michelson-Morley Experiment, and it is not going to materially affect train travel where one deals with velocities of the order of a few tens of miles per hour. But, obviously, when man has conquered space and travels amongst the stars at high velocities the effect is likely to be appreciable. Distances which from Earth appear enormous will appear much less to our rapidly moving space-man; so much shorter that journeys which appear impossible to complete within a single lifetime to his Earthbound cousin may be possible. However, we cannot consider this spatial contraction in isolation. Let us consider the second prediction arising from the Lorentz transformation. Suppose we set out to measure the time that a journey takes, and we leave the point x_1 in the station when the station clock reads a time t_1 and pass the point on the track x_2 at a time t_2 according to the station clock, then the station master would say that the time T of the journey was

$$T=t_2-t_1. \tag{3.23}$$

However, suppose that you decide to time the journey by your wrist watch and that you leave the station at time t'_1 and pass the point x_2 at time t'_2. Then you will claim that the time taken for the journey is T' where

$$\begin{aligned} T' &= t'_2-t'_1 \\ &= \frac{t_2-ux_2/c^2}{\sqrt{(1-u^2/c^2)}}-\frac{t_1-ux_2/c^2}{\sqrt{(1-u^2/c^2)}} \\ &= \frac{t_2-t_1}{\sqrt{(1-u^2/c^2)}}=\frac{T}{\sqrt{(1-u^2/c^2)}}. \end{aligned} \tag{3.24}$$

Thus you would say that the journey did not take as long as the station master said it had. In fact, if the train travelled with the velocity of light, you would say that the journey had been instantaneous. Of course, this descrepancy is not surprising, since we would have argued about the length of the journey in any case.

Again, this is a negligible effect on Earthbound trains. But suppose there are two young identical twins and one stays on Earth while the other makes a long journey in space at a high velocity. When the space-ship returns, the twin who has stayed on Earth will be old and grey, while the one who made the journey will still only be in middle age. In fact, were we to contemplate a round journey to any but the very nearest stars we would not return to Earth until all our contemporaries had been dead for many thousands of years.

Naturally, the addition law for velocities (3.4) will also require to be changed. Now

$$\begin{aligned} V'_x &= \frac{x'_2-x'_1}{t'_2-t'_1}=\frac{[x_2-x_1-u(t_2-t_1)]/\sqrt{(1-u^2/c^2)}}{[t_2-t_1-u(x_2-x_1)/c^2]/\sqrt{(1-u^2/c^2)}} \\ &= \frac{V_x-u}{1-uV_x/c^2}, \end{aligned} \tag{3.25}$$

while

$$\begin{aligned} V'_y &= \frac{y'_2-y'_1}{t'_2-t'_1}=\frac{y_2-y_1}{[(t_2-t_1)-u(x_2-x_1)/c^2]/\sqrt{(1-u^2/c^2)}} \\ &= \frac{V_y\sqrt{(1-u^2/c^2)}}{1-uV_x/c^2}, \end{aligned} \tag{3.26}$$

and similarly

$$V'_z=\frac{V_z\sqrt{(1-u^2/c^2)}}{1-uV_x/c}. \tag{3.27}$$

Again we notice that as u/c and V_x/c become very small, these equations reduce to the classical equations (3.4).

So much for the changes in kinematical measurements. What happens to our concepts in mechanics, i.e. what happens to relative measurements of quantities like mass and energy?

To obtain the new laws of mechanics we turn to the first of Einstein's postulates, namely that the laws of physics will be the same for all inertial frames. In particular, we require that the law of conservation of momentum will be the same in all frames. This is only possible provided we are willing to allow the mass of a body to be a function of its velocity.

Now let us consider the following experiment. Two observers O_1 and O_2 in Fig. 14 are moving with velocity v relative to one another parallel and antiparallel to the x-axis respectively. Both observers have identical billiard balls which have the same mass m_0 when at rest.

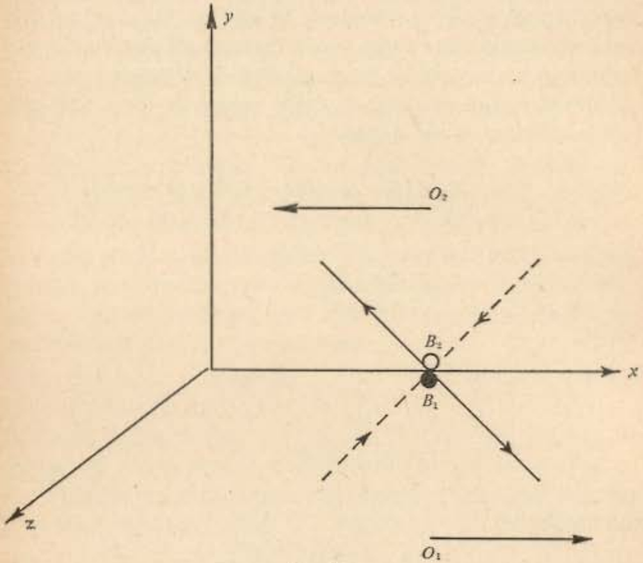


Fig. 14.

Each of the observers at the same time throws his ball at the other with a velocity V, which from his own point of view is perpendicular to the x-axis. The balls will, of course, collide and each will rebound to the observer who threw it. Let us look at what happens from the point of view of the observer O_1 (Fig. 15).

The y-component of the velocity of the ball B_1 will be initially V, and after the collision $-V$. Both before and after the collision the velocity in the x-direction will be zero. The second ball, B_2, will appear to have a constant velocity v in the x-direction and a velocity $-V\sqrt{(1-v^2/c^2)}$ in the y-direction before the collision, and $V\sqrt{(1-v^2/c^2)}$ in the y-direction after the collision. Thus the total momentum before the collision in the y-direction is

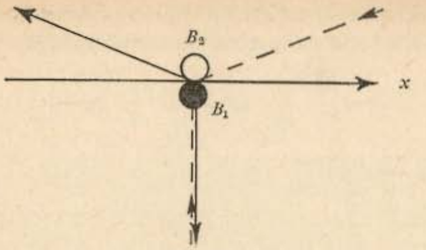


Fig. 15.

$$m(V)V - m(\sqrt{[V\sqrt{(1-v^2/c^2)}^2+v^2]})V\sqrt{(1-v^2/c^2)}$$

and this must be equal to the total momentum in the y-direction after the collision, which is

$$-m(-V)V + m(\sqrt{[-V\sqrt{(1-v^2/c^2)}^2+v^2]})V\sqrt{(1-v^2/c^2)}.$$

Thus

$$m(V)+m(-V)=2\sqrt{(1-v^2/c^2)}m(\sqrt{[V^2(1-v^2/c^2)+v^2]}). \quad (3.28)$$

Suppose we now set $V=0$, then

$$m(v)=\frac{m_0}{\sqrt{(1-v^2/c^2)}}. \quad (3.29)$$

Now we see the third of the startling predictions of the Special Theory of Relativity. Not only do lengths and times contract the faster the observer travels but also an object becomes heavier the faster it travels, until at the velocity of light it is going everywhere infinitely quickly, it is infinitely short and has acquired an infinite mass.

Let us continue our discussion of relativistic mechanics and consider what happens to our ideas of energy.

Suppose we have an object of rest mass m_0 sitting at the origin of our reference frame at time $t=0$ and we apply a constant force F to it until it reaches the point x_1, by which time it will have accelerated up to some velocity v_1. By definition the total work done on the particle is

$$W=\int_0^{x_1} Fdx=\int_0^{x_1}\frac{dp}{dt}\,.\,dx=\int_0^{x_1}\frac{dx}{dt}\,.\,dp=\int_0^{x_1} vdp \qquad (3.30)$$

where p is the momentum mv of the particle; we have used Newton's law of motion which equates the force F to the rate of change of momentum dp/dt and the fact that dx/dt is the velocity v of the particle. Integrating* (3.30) we have that

$$W=m_0c^2[-1+(1-v_1^2/c^2)^{-\frac{1}{2}}]. \qquad (3.31)$$

Now, by the law of conservation of energy, we see that this must be equal to the gain in kinetic energy, and certainly for v_1/c very small we see that (3.31) reduces to

$$W=\tfrac{1}{2}m_0v_1^2$$

as we should expect in classical physics.

We see that in general W can be written as two terms

$$\begin{aligned} W&=E(v)-E(0)\\ &=\frac{m_0c^2}{\sqrt{(1-v^2/c^2)}}-m_0c^2. \end{aligned} \qquad (3.32)$$

The energy that the particle has due to its motion is obviously $E(v)$. The energy $E(0)$ in no ways depends on the velocity of the particle. We can, in fact, write (3.32) as

$$E(v)=W+E(0).$$

Thus $E(v)$ is the total energy of the particle, being the sum of the work done, W, by the applied force in giving the particle the velocity v plus the term $E(0)$, which is an intrinsic energy associated with the fact that the particle has a non-zero rest mass

$$\begin{aligned} E(0)&=m_0c^2\\ E(v)&=m(v)c^2. \end{aligned} \qquad (3.33)$$

Not only does the particle possess an energy $E(0)$ when it is at rest, but what an enormous energy it is! We see that one gramme of material is equivalent to an energy of 9×10^{20} ergs, or enough energy to produce a million kilowatts of electric power for 24 hours. It was the release of a small fraction of this energy which destroyed Hiroshima and Nagasaki.

We quote one more equation for future reference, and that relates the energy $E(v)$ of a particle to its momentum $p=mv$

$$E^2=m^2c^4=c^2p^2+m_0^2c^4. \qquad (3.34)$$

One extraordinary result should be mentioned. While the laws of mechanics have had to be seriously modified as a result of the study of relativity, the laws of electromagnetism as given by Maxwell remain unchanged (Appendix A). Not only do they remain unchanged but they take on a new beauty and simplicity in a relativistic formulation.

We close this chapter by merely stating that the predictions of the Special Theory of Relativity have been experimentally verified on countless occasions. The relativistic mass-energy relationship, the Lorentz contraction and the relativistic time dilation are all seen in action every day in modern atomic reactors and particle accelerators.

So far we have only considered uniform motion and have ignored the possibility of accelerations. The study of non-uniform motion and its relationship with gravitation forms the subject of Einstein's General Theory of Relativity. A detailed study of general relativity is beyond the scope of this small volume. However, we shall return in the last chapter to discuss a few ideas contained in the general theory.

---

* Readers unfamiliar with integral calculus are asked to accept the result (3.31), the proof of which we give below

$$p=mv=m_0v(1-v^2/c^2)^{-\frac{1}{2}} \text{ by equation (3.29)}$$

Thus $$dp/dv=m_0[(1-v^2/c^2)^{-\frac{1}{2}}+v^2/c^2(1-v^2/c^2)^{-3/2}]$$

and $$\begin{aligned}\int_0^{p_1} vdp&=\int_0^{v_1} vm_0[(1-v^2/c^2)^{-\frac{1}{2}}+v^2/c^2(1-v^2/c)^{-3/2}]dv\\ &=m_0c^2[(1-v_1^2/c^2)^{\frac{1}{2}}-1]\\ &\quad+m_0/2c^2[2v_1^2c^2(1-v_1^2/c^2)^{-}\ +4c^4(1-v_1^2/c^2)^{\frac{1}{2}}-4c^4]\\ &=-m_0c^2+m_0c^2(1-v_1^2/c^2)^{-\frac{1}{2}}\end{aligned}$$

4. The Interaction of the Electromagnetic Field with Matter

Just as the Michelson-Morley Experiment had required a radical revision of classical concepts for its explanation, there were a number of other empirical observations which were impossible to explain in terms of the kinetic theory of matter and Maxwell's electromagnetic theory alone. In this chapter we shall examine a few of these experimental results and discuss the attempts and failures to give them a classical explanation.

LINE SPECTRA

At the very beginning of the nineteenth century Fraunhofer had set out to make a detailed study of the spectrum of sunlight and had discovered that instead of being formed from a continuous band of slowly varying colours, as it had appeared at first to Newton, there existed in addition a number of very fine, but clear and distinct, dark lines. In other words, it appeared that there were very sharply defined wavelengths of light which do not reach us from the Sun.

By the second half of the nineteenth century Kirchhoff and Bunsen had laid the foundations of spectral analysis based on the observation that the vapour of an element when heated until it glows does not radiate randomly but only at very specific wavelengths. Thus instead of emitting a continuous band of colour a gaseous element will only radiate very sharp lines of colour, hence the expression line spectra. These lines, or wavelengths, are found to be characteristic of the element forming the vapour, and this can be seen, for example, in the characteristic yellow light of sodium street lamps.

Not only will an element emit only characteristic wavelengths but it was also discovered that the vapour will only absorb light of specific wavelengths. And, finally, these absorption and emission wavelengths are the same for any given element.

Using this result, we can at once give an explanation of the Fraunhofer dark lines. Suppose that in the extremely hot interior of the Sun all possible wavelengths of radiation are generated; then, as the light passes through the relatively cool corona of the Sun, those wavelengths corresponding to the characteristic absorption lines of the elements in the corona will be absorbed, and light of these wavelengths will not reach us from the Sun. In this way we can predict what gases are present in the Sun's corona, or, for that matter, the corona of any star. For example, it was from such a study of the Fraunhofer dark lines that the gas helium was found to exist in the Sun's corona before it was discovered on Earth.

However, we are still left with the problem of explaining the origin of these characteristic emission and absorption wavelengths in vaporised elements. As we have already suggested, the source of the emitted radiation is the motion of the charged ions in the material. But why is this motion not random? Why do the charges appear to vibrate in a very special manner characteristic of the element involved?

Let us now leave the subject of the emission and the absorption of light by vaporised elements and consider how thermal radiation is emitted and absorbed by a solid.

In 1879 Stefan discovered the empirical law which bears his name. The rate of emission of thermal radiation R per unit area of the surface of a body is proportional to a constant e between 0 and 1 which depends on the nature of the emitting surface, and is called the emissivity of the surface, and to the fourth power of the temperature T.

$$R = \sigma e T^4 \tag{4.1}$$

where σ is a new physical constant, the Stefan-Boltzmann constant, equal to $0{\cdot}567 \times 10^4$ erg cm^{-2} deg^{-4} s^{-1}.

Suppose that instead of studying the emission of thermal radiation by a surface we now shine thermal radiation on to a surface and examine how much of it is absorbed. The ratio of the amount absorbed to the amount incident we shall call the absorptivity a. In 1895 Kirchhoff established the following simple relationship

$$e=a. \tag{4.2}$$

BLACK-BODY RADIATION

By a *black body* we mean one which will absorb all radiation incident on it, i.e. $a=1$. (This should be compared with a vaporised element which will only absorb a very small fraction of the radiation incident on it, corresponding to the associated characteristic wavelengths.)

By Kirchhoff's relation (4.2) a black body is not only the most efficient absorber of radiation but also the most efficient emitter of radiation.

If we take a black body and shine thermal radiation on to it at a constant rate, it will start to heat up and its temperature will rise, and as its temperature rises it will emit radiation at a greater rate according to Stefan's Law (4.1). The temperature will continue to rise until it reaches a value where the rate of emission of thermal radiation is exactly equal to the rate of absorption of radiation, at which point we say that the system is in thermodynamic equilibrium and has reached a steady state.

If we study the spectral distribution $R(\lambda)$ of the black-body radiation in the steady state, then we obtain the typical curves shown in Fig. 16 by the solid lines.

Wien in 1893, from very general arguments, was able to prove the following relationship

$$R(\lambda)=f(\lambda T)/\lambda^5 \tag{4.3}$$

where $f(\lambda T)$ is an unknown combination of constants such as the Boltzmann constant or the velocity of light, etc., together with the wavelength λ and the temperature T. But whatever the combination, λ and T always enter it as the product λT.

Rayleigh and Jeans then set out to give an explanation of these results in terms of the kinetic theory of matter, using Boltzmann statistics (Appendix B). They considered a model in which the ions oscillate with a set of continuous frequencies about their mean positions. The result of their classical analysis is

$$R(\lambda)=(8\pi k/\lambda^5)\lambda T. \tag{4.4}$$

This is seen to be in agreement with Wien's Law, where $f(\lambda T)=8\pi k\lambda T$. However, as can be seen from the broken line in Fig. 16, it is in complete disagreement with experiment except at very long wavelengths. This violent disagreement with experiment is known as the ultra-violet catastrophe, since the disagreement becomes particularly marked at wavelengths in the ultra-violet region and below.

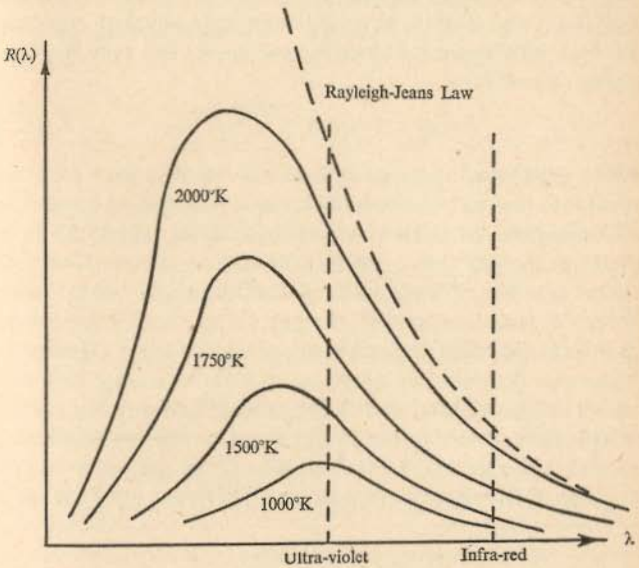


Fig. 16.

The ultra-violet catastrophe was the source of much consternation amongst physicists at the turn of the century and was finally resolved by Planck in 1901. However, Planck's solution of the problem involved the introduction of a new postulate, which was to lead to a revolution in our understanding of the physical world at least as great as that caused by relativity.

Planck postulated that the oscillating ions which were supposed to be the source of the emitted black-body radiation could not emit energy continuously – as they must if they were classical oscillators – but could only emit energy in multiples of a basic unit $h\nu$ where ν is the frequency of the oscillator and h is a new, fundamental, physical constant $h = 6{\cdot}6 \times 10^{-27}$ erg s^{-1}. Hence the oscillators can only have an energy E satisfying

$$E = nh\nu \qquad n = 0, 1, 2, 3, \ldots . \tag{4.5}$$

When a physical quantity is built out of a number of basic units it is said to be *quantised.* A system containing a number of basic units of some physical quantity is said to be in a *quantum state* of that quantity with a *quantum number* equal to the number of basic units in that quantum state. Hence Planck's postulate is that energy of the oscillating ions is quantised and that the quantum of energy is $h\nu$. The oscillators can only exist in quantum states with energy $E_n = nh\nu$ which can be labelled by the energy quantum number n.

This discreteness in the energy spectrum of the oscillators requires an alteration in the statistics which we apply in our analysis of the problem (Appendix B). The result of Planck's analysis is that

$$R(\lambda) = \frac{8\pi hc}{\lambda^5} \; \frac{1}{\exp(hc/k\lambda T) - 1}. \tag{4.6}$$

We see at once that this agrees with Wien's Law where

$$f(\lambda T) = \frac{8\pi hc}{\exp(hc/k\lambda T) - 1}, \tag{4.7}$$

and that at large values of λ (4.6) reduces to the Rayleigh-Jeans Law (4.4). But, most important of all, it is in excellent agreement with experiment over a wide range of temperatures and wavelengths.

For a few years Planck's postulate existed in a vacuum. Certainly it explained black-body radiation, but it appeared as a totally *ad hoc* assumption designed only to explain this single class of observations. Even Planck could not bring himself to believe in the quantisation of the ionic oscillators and was continually looking for a way round this assumption, without success. Then, in 1905, the same year that he published his special theory of relativity, Einstein showed how a simple generalisation of Planck's postulate could explain another strange phenomenon, the photo-electric effect.

THE PHOTO-ELECTRIC EFFECT

As mentioned previously, Hertz had discovered experimentally the existence of the electromagnetic waves predicted by Maxwell's theory. Using an oscillating circuit containing a spark gap he discovered that he could produce a spark in a second, suitably tuned, circuit. But, more than this, he discovered that the spark in the second circuit was more readily produced if light was shone on to the electrodes of this circuit.

Shortly after the announcement of Hertz's discovery, Hallwachs showed that when many metals were illuminated with ultra-violet light they gained a small positive electric charge. There were two possible explanations for this observation, either the light deposited some positive charge on the metal, or it caused the flow of some negative charge out of the metal.

Stoletow devised a system which showed conclusively that in fact some negative charge left the metal. He constructed a circuit (Fig. 17) in which a photo-electrically sensitive plate is connected to the negative terminal of a battery. Another non-sensitive plate was connected to the positive terminal through a sensitive galvonometer. When ultra-violet light was shone on to the sensitive plate, a current flowed in the circuit. However, if the terminals were reversed no current flowed.

The question which immediately arises is, 'What is it that carries the negative electric charge across the gap between the plates?' The current persists even if the system is put in a vacuum, so that it cannot be carried by the air molecules. Of course, from electrolysis we know that within the electrodes there can exist positive and negative ions, perhaps they carry the charge. Lenard removed this possibility with the following ingenious experiment. He used as the sensitive plate in Stoletow's circuit a sodium amalgam, and for the non-sensitive plate a clean platinum wire. Now, if the current was carried by the sodium ions, then after some time sodium deposits would accumulate on the platinum; no sodium deposits were discovered.

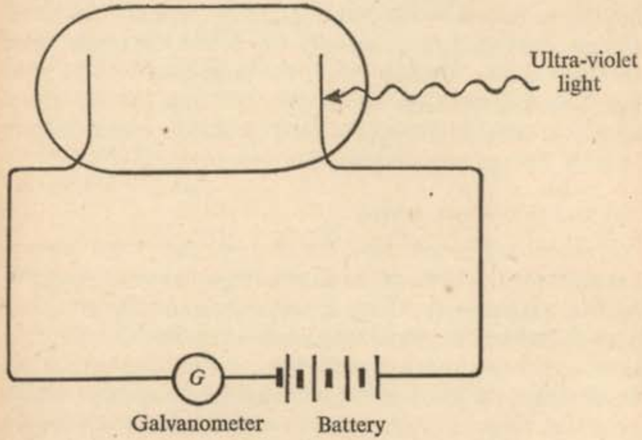


Fig. 17.

This mystery was resolved in 1897 with the discovery by J. J. Thomson of the *electron.*

If two electrodes enclosed in a glass tube (Fig. 18) which is pumped down to an extremely low pressure (approximately 0·01 mmHg) are connected to the terminals of a high-voltage battery, then a current will flow in the circuit and a luminous spot appears on the glass behind the positive electrode. If a screen is placed in front of the negative electrode (or cathode, as it is usually called), the spot disappears. If an electric field is applied to the glass tube perpendicular to its length, the spot will move to a position S', i.e. towards the higher voltage

region. This suggests that whatever it is that causes the spot comes from the cathode and carries a negative electric charge. This stream of negative electricity is now called *cathode ray,* the apparatus in Fig. 18 is called a *cathode ray tube* and is, of course, the basis of modern television.

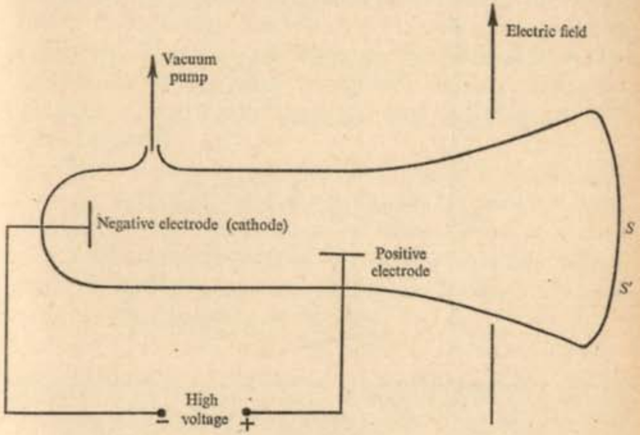


Fig. 18.

A more detailed study of the motion of the luminous spot under the action of electric and magnetic fields supports the view that the cathode ray is a stream of negatively charged particles which we now call electrons.

The next point established was that the electrons were always the same, no matter from what material they came. We should also note in passing that since the electrons carry a negative electric charge and are emitted from metals which are originally electrically neutral, there must reside within metals a source of positive charge. This was already known from electrolysis, where the existence of positive ions had been established. However, we shall postpone a detailed discussion of these points until the next section.

We have seen that the emission of electrons from a metal is facilitated by the action of light shining upon it. Lenard, in a series of experiments starting in 1900, set out to study this photo-electric effect in some detail. Starting with a circuit essentially the same as that indicated in Fig. 17, if we plot the current which flows in the galvonometer against the applied voltage we obtain a curve of the form shown in Fig. 19. Not only does the current flow more easily when light is shone on to the cathode but it even requires a voltage V_0 to be applied in the opposite direction in order to stop the current flowing.

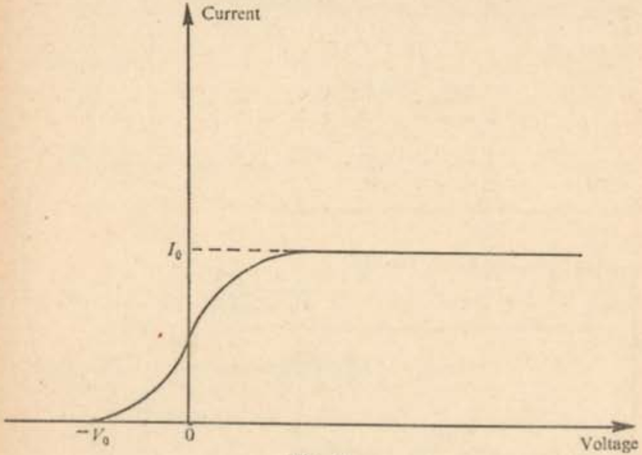


Fig. 19.

The immediate answer to this observation, from a classical point of view, is that the energy carried by the light is used to break the bonds binding the electrons into the metal – or, at least, weaken them – so that when an electric field is applied they are emitted more easily. However, if this is so, then by increasing the intensity of the light, i.e. the amount of light and hence the energy, falling on to the metal, we should be able to release electrons even more easily; that is, V_0 should become larger. This is not found to happen, provided the frequency of the light is kept constant; no matter how much we increase the intensity, V_0 remains the same. For any given voltage, as we increase the intensity of the light the current increases. On the other hand, if we change the frequency of the light, V_0 changes immediately, V_0 increasing as the frequency increases. No matter how low we make the intensity of the light, electrons are still given off, provided the frequency is high enough. The critical frequency, below which V_0 is zero, depends entirely on the nature of the metal surface.

There is obviously a paradox here. For, if light is a wave, its energy should be spread all over the wave front, and hence an extremely localised object like an electron should take a long time to acquire enough energy to be released from the metal when the intensity is extremely low. However, electrons are emitted almost instantaneously when light strikes the surface of a metal.

Einstein's solution to this problem was both simple and baffling. He suggested that light, in this context, did not have a wave nature but more closely resembled Newton's corpuscles. Each corpuscle of light, or *photon*, as it is now called, he assumed carried a quantum of energy E which was proportional to the frequency of the light

$$E = h\nu \tag{4.8}$$

the constant of proportionality h is again Planck's constant. In this way it is clear that by reducing the intensity, i.e. the number of photons incident per second, we only reduce the number of electrons emitted and hence the current. But as long as the frequency is high enough, so that the energy $h\nu$ is greater than the binding energy W of the electrons, electrons will be emitted. The maximum kinetic energy E_{max} of the electrons emitted should be

$$E_{\text{max}} = h\nu - W \tag{4.9}$$

and in order to stop the current flowing we must absorb this kinetic energy by an applied potential, i.e. $V_0 = E_{\text{max}}$.

In 1916 Millikan tested this law experimentally and found excellent agreement with the theory. The value of h in this case was found to agree with that used by Planck to explain black-body radiation to better than 0·5 per cent.

RUTHERFORD'S NUCLEAR ATOM

Between 1896 and 1898, Becquerel and Mme Curie had begun the investigation of the phenomenon known as natural radioactivity. That is, the observation that several of the heaviest elements in nature, in particular thorium and uranium, emit penetrating radiation which can be detected by its capacity to blacken a photographic plate. On closer examination this radiation is found to have three very distinct components. These have been named α-, β- and γ-rays.

The α-rays are found to come to rest after passing through only a few centimetres of air. They have the greatest mass of the three rays, being made up of particles with a mass approximately the same as the helium atom, or four times the mass of the hydrogen atom. They are also found to possess a positive electric charge which is exactly twice the magnitude of the charge on the electron.

The β-rays are more penetrative than the α-rays and are capable of passing through thin sheets of metal foil. They, in fact, turn out to be identical with Thomson's cathode rays. The electrons are much lighter than the α-particles, having a mass of approximately 1/2000 of the mass of a hydrogen atom.

The γ-rays are by far the most penetrative of the three, being able to pass easily through several centimetres of the lighter elements, e.g. aluminium. The γ-particles are found to possess no electric charge and no rest mass, they are, in fact, simply extremely high-frequency electromagnetic radiation. γ-radiation has now become the generic name for all electromagnetic radiation at frequencies beyond the ultra-violet region. However, at wavelengths around 10^{-8} cm this radiation is more commonly called X-rays.

The positively charged α-particles and the negatively charged electrons must obviously, in most cases, be bound together in some stable configuration to form the neutral atoms of which matter is composed. What is this configuration?

J. J. Thomson proposed a tentative model in which the atom was composed of a uniform, spherical distribution of positive charge in which were imbedded the electrons. The actual position of the electrons was to be determined by conditions of electrostatic stability. At low temperatures the electrons would be at rest in these equilibrium positions. On heating, however, they would become excited and vibrate about these positions. Since the electromagnetic theory predicts that accelerated charged particles, such as the vibrating electrons, emit electromagnetic radiation, we should hope that the allowed modes of vibration of the electrons would give rise to radiation in agreement with the observed line spectra of the elements. Unfortunately, detailed calculations were unable to predict spectra in agreement with the experimental observations.

In 1911 Rutherford carried out a classic experiment which showed the inadequacy of the Thomson model. Rutherford fired the newly discovered α-particles at a thin metal foil and then measured the angle through which they were scattered (Fig. 20).

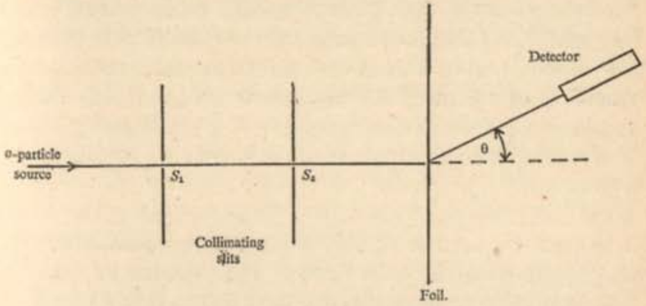


Fig. 20.

Since in the Thomson model the electrons are uniformly distributed throughout the positive charge region, the effect of the charges will be to cancel each other out. Hence a positively charged α-particle approaching such an atom would not see any concentrated region of positive charges, and thus at no time would it experience any extremely large electrostatic repulsions. A detailed calculation based on these assumptions suggests that the majority of the α-particles will not be scattered through large angles, i.e. in Fig. 20 nearly all the α-particles would be expected to be observed at angles of θ less than 5°. This was not the result that Rutherford found. Not only were many α-particles scattered through angles much larger than those predicted by the Thomson model, some were actually back scattered, i.e. θ was greater than 90°.

In order to explain these observations, Rutherford assumed that nearly all the mass and all the positive charge was concentrated at the centre of the atom in what he called the *nucleus*. Around this nucleus there was a cloud of electrons. This model gave brilliant agreement between the predicted and the experimentally observed angular distribution of the scattered α-particles. However, it was quickly shown by Ehrenfest that such a model could never be stable. Certainly, the electrons could orbit the nucleus in the same way as the planets orbit the Sun – after all, the electric force law is identical in form with that for gravity – but, being charged particles moving in accelerated orbits, the electrons would necessarily have to emit electromagnetic radiation and hence lose energy. As they lost energy they would spiral in towards the nucleus, and in a short time all the atoms in the universe would have collapsed. In any case, the prediction of the continuous emission of radiation is not in agreement with the observed discrete spectrum which is known to be emitted by free atoms.

In conclusion, while we may seem to have explained several phenomena this has only been at the price of introducing fresh confusion to the scene. We have had to give up some of our classically based beliefs. Light, which until now we have considered to be a wave process, now takes on particle aspects when we discuss its emission or absorption. The only model of the atom which will agree with the α-particle scattering observations appears to be electromagnetically unstable. And we still have no explanation of line spectra.

5. The Old Quantum Theory

THE BOHR ATOM

In 1911 Bohr came to Manchester to work with Rutherford. Faced with the baffling contradiction between Rutherford's experimental results confirming the nuclear structure of the atom and the classical observation that such a structure must be unstable, Bohr put forward new postulates which explained the structure of the atom and at the same time accounted for the line spectra of the elements. These postulates were:

1. The electrons move in circular orbits around the nucleus under the influence of the electromagnetic attraction and obeying the classical laws of electromagnetism.
2. Instead of the infinity of possible orbits which are allowed by classical mechanics, only those with an orbital angular momentum which is an integral multiple of Planck's constant h divided by 2π are allowed.
3. Even although the electron is accelerated, it does not radiate as long as it remains in one of these allowed orbits.
4. If the electron moves from an allowed orbit of energy E_1 to another of lower energy E_2, then it will radiate. The frequency ν of the emitted electromagnetic radiation will be equal to the difference in energy $(E_1 - E_2)$ divided by Planck's constant h.

We see that in addition to the quantisation of energy introduced by Planck and Einstein, Bohr's postulates require the concept of the quantisation of angular momentum.

Let us examine the structure of a one-electron atom from the point of view of these postulates. We consider a single electron with charge $-e$ orbiting a nucleus of charge $+Ze$ (see Fig. 21).

For equilibrium, the electrostatic attraction of the electron to the nucleus Ze^2/r^2 must be exactly equal to the centrifugal force mv^2/r.

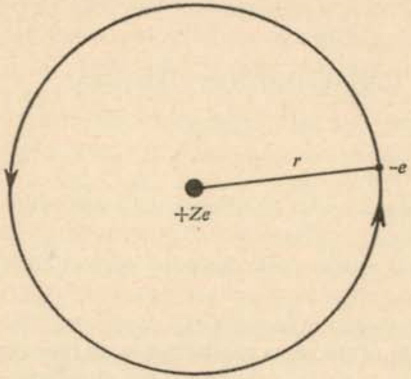


Fig. 21.

The angular momentum L for a circular orbit is defined to be the product of the radius of the orbit times the linear momentum of the particle. By the second postulate this must be quantised, hence

$$L = r \times mv = \frac{nh}{2\pi} \qquad (5.1)$$

where n is an integer 1, 2, 3, ...

By the balancing of the electrostatic and the centrifugal force we have

$$Ze^2 = mv^2 r = \frac{n^2h^2}{4\pi^2 mr}. \qquad (5.2)$$

Hence the radii of the allowed orbits are given by

$$r = \frac{n^2h^2}{4\pi^2 mZe^2} \qquad n = 1, 2, \ldots . \qquad (5.3)$$

Now the electromagnetic energy plus the kinetic energy of the electron is (using (5.2))

$$E_n = -\frac{Ze^2}{r_n} + \tfrac{1}{2}mv_n^2 = -\tfrac{1}{2}\frac{Ze^2}{r_n}$$

$$E_n = -\frac{4\pi^2 mZ^2e^4}{2h^2n^2}. \qquad (5.4)$$

SPECTRAL LINES AND ATOMIC STRUCTURE

Bohr's fourth postulate, coupled with the result (5.4), gives us for the first time a discrete set of spectral lines. If an electron were to move from an orbit of angular momentum $n_1h/2\pi$ to another of smaller angular momentum $n_2h/2\pi$, it would emit electromagnetic radiation of frequency

$$\nu_{12} = \frac{2\pi mZ^2e^4}{2h^3}\left(\frac{1}{n_2^2} - \frac{1}{n_1^2}\right). \qquad (5.5)$$

Many years previously, physicists studying the experimentally observed spectra of hydrogen had discovered some startling regularities in the wavelengths of various lines, which were then arranged in series and named after their discoverers. They were:

(i) the Lyman series in the ultra-violet region with wavelengths λ given by

$$\frac{1}{\lambda} = R_H\left(1 - \frac{1}{n^2}\right), \qquad n = 2, 3, 4, \ldots;$$

(ii) the Balmer series in the near ultra-violet and the visible region with wavelengths given by

$$\frac{1}{\lambda} = R_H\left(\frac{1}{4} - \frac{1}{n^2}\right) \qquad n = 3, 4, 5, \ldots;$$

(iii) the Paschen series in the infra-red region with

$$\frac{1}{\lambda} = R_H\left(\frac{1}{9} - \frac{1}{n^2}\right) \qquad n = 4, 5, 6, \ldots;$$

(iv) the Brackett series in the infra-red region with

$$\frac{1}{\lambda} = R_H\left(\frac{1}{16} - \frac{1}{n^2}\right) \qquad n = 5, 6, 7, \ldots;$$

(v) the Pfund series, also in the infra-red region, with

$$\frac{1}{\lambda}=R_H\left(\frac{1}{25}-\frac{1}{n^2}\right) \qquad n=6, 7, 8, \ldots$$

where R_H, the Rydberg constant, has the measured value 109 677·576±0·017 cm^{-1}. These, we now see, are simply special cases of Bohr's general result (5.5). According to Bohr's theory

$$R_H=\frac{2\pi me^4}{2h^3c} \tag{5.6}$$

since $Z=1$ for the hydrogen atom and $\lambda\nu=c$. However, (5.6) is not quite correct, because we assumed in its derivation that the nucleus was at rest with the electron orbiting it. In fact, the *proton*, which is the name given to the nucleus of the hydrogen atom, and the electron will both rotate about their mutual centre of mass. This leads to m, the mass of the electron in (5.6), being replaced by $mM/(m+M)$ where M is the mass of the proton. Since the proton mass is approximately 2000 times the electron mass, this is a small correction. But when the values of all the quantities entering the right-hand side of equation (5.6) are now inserted, we obtain a

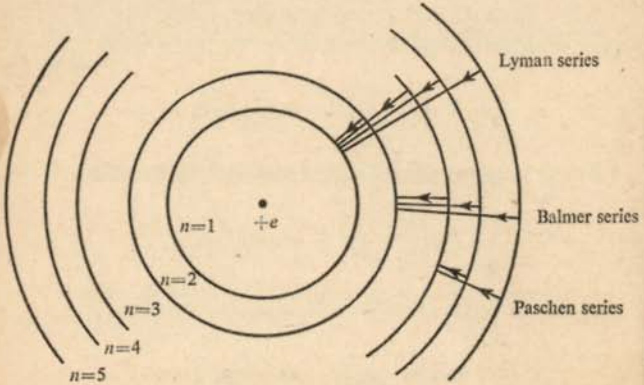


Fig. 22.

predicted value of R_H of 109 681 cm^{-1}, i.e. an agreement with experiment to three parts in 100 000.

We have seen that through the quantisation of the angular momentum we have obtained a quantisation of the energy levels of the hydrogen atom with the resultant spectral lines (see Fig. 22), and this theory is in excellent agreement with experiment. Certain other elements are also fairly well explained by this simple theory, namely the alkali metals sodium, potassium, etc., which behave as if they contained a single electron orbiting a shielded core. However, for other elements the theory is not in very good agreement with experiment and naturally becomes extremely complicated when the interaction between the electrons has to be taken into account.

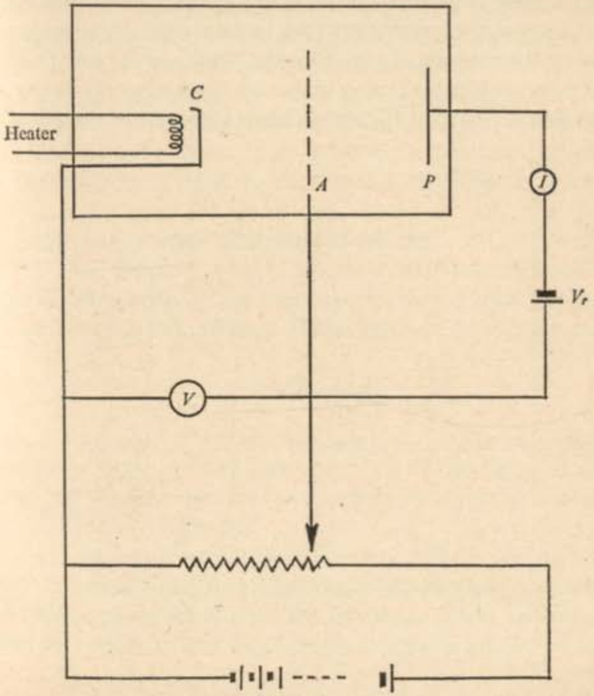


Fig. 23.

But, none the less, the general statement that the energy levels of all atoms are quantised, and that characteristic spectra of the elements come from the transitions of the electrons from one allowed orbit to the next, remains true.

The empirical verification that the energy levels of atoms are quantised was produced by Franck and Hertz in 1914. Their apparatus is sketched in Fig. 23. Electrons are emitted from the cathode C by heating it. They are then accelerated towards the anode A under an applied potential V. The anode is made in the form of a grid so that some of the electrons may pass through it. Behind the anode there is a plate P, between it and the anode there is a small retarding potential V_r. If the electrons have enough kinetic energy on passing the anode to overcome this retarding potential, they will reach the plate and a current I will flow in the circuit. If the temperature of the cathode is such that the electrons are just emitted but have no kinetic energy, then any kinetic energy they may possess on reaching the anode must come from their loss in potential energy in moving from C to A, this having a maximum value of eV. A plot of the current I is made against the applied voltage V, resulting in a curve similar to that sketched in Fig. 24.

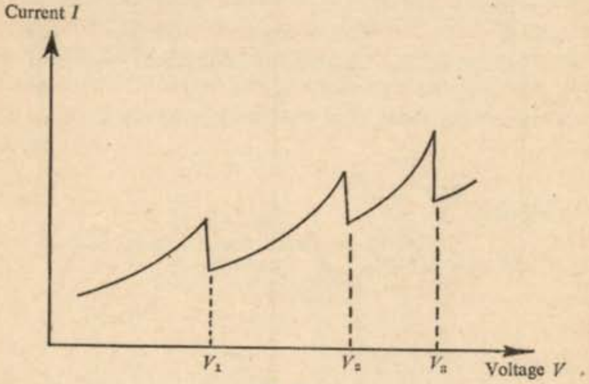


Fig. 24.

We see that at a series of discrete voltages there is a sudden sharp drop in the current, i.e. in the number of electrons reaching the plate per second. Thus we may assume that a substantial number of the electrons have lost some of their kinetic energy to the atoms of the gas between the cathode and the anode, and hence do not have enough kinetic energy left to overcome the small retarding potential V_r. Amongst other things we see that the atoms can only receive their energy in discrete amounts corresponding to the potentials $V_1, V_2, V_3, \ldots$ in Fig. 24. Moreover, these potentials are found to agree very accurately with the discrete absorption spectra of vaporised elements.

THE WILSON-SOMMERFELD QUANTISATION RULES*

The work of Planck, Einstein and Bohr had certainly given striking agreement between theory and experiment for black-body radiation, the photo-electric effect and the spectral lines of hydrogen, although the bases of the postulates that had been required for these new theories were still completely obscure. In 1916 Wilson and Sommerfeld carried the story one stage further, when they enunciated the following general rule for the quantisation of any system in which the co-ordinates are periodic functions of the time.

For any physical system in which the co-ordinates are periodic functions of the time, there exists a quantum condition for each co-ordinate. These quantum conditions are

$$\oint p_q \, dq = n_q h$$

where q is one of the co-ordinates, p_q is the momentum associated with that co-ordinate, n_q is a quantum number taking on integral values and $\oint$ means integration over one period of the co-ordinate q.

This rule includes as special cases both Planck's postulate for black-body radiation and Bohr's postulate of angular momentum quantisation of the hydrogen atom. In Planck's case, the position of an oscillating body vibrating along the

---

* This section may be omitted by readers who are not familiar with integral calculus.

x-axis can be expressed in the form $x = x_0 \cos \omega t$ where $\omega = 2\pi\nu$. Its velocity is then given by $V = -\omega x_0 \sin \omega t$. The Wilson-Sommerfeld rule then becomes

$$\oint p_x \, dx = -\oint m\omega x_0 \sin \omega t \, dx = \int_0^{2\pi/\omega} m\omega^2 x_0^2 \sin^2 \omega t \, dt$$

$$= m x_0^2 \, \omega\pi = nh. \tag{5.7}$$

The energy of such an oscillator is $E = \frac{1}{2} m x_0^2 \omega^2$, which from (5.7) gives

$$E = nh\nu \qquad n = 1, 2, 3, \ldots, \tag{5.8}$$

i.e. the energy of the oscillators is quantised as postulated by Planck. If we consider Bohr's model of the hydrogen atom and take as our co-ordinate the azimuthal angle of the electron in its orbit, the rule becomes

$$\oint p_\phi \, d\phi = \int_0^{2\pi} L d\phi = \int_0^{2\pi} mvr \, d\phi = mvr 2\pi = nh$$

$$L = \frac{nh}{2\pi} \tag{5.9}$$

which is simply a statement of Bohr's second postulate.

If one examines, with very high resolution, the spectral lines of hydrogen, they are found not to be single lines but groups of lines, the spacing between members of the group being some 10^{-4} times the spacing between the groups. Sommerfeld gave the following explanation of this hyperfine splitting of the spectral lines. The electromagnetic law of force, like the gravitational force, implies not that the electrons must move in circular orbits but only that they must move in elliptic orbits (a circle is a limiting case of an ellipse). There are then two quantisation rules, one governing the azimuthal angle as before, and another governing the distance of the electron from the nucleus, which was a constant in Bohr's circular orbits. In this way, Sommerfeld obtained two quantum numbers, the principal quantum number n (Bohr's quantum number) and the azimuthal quantum number n_ϕ, which defines the eccentricity of the ellipse. Now, if an electron moves in an elliptic orbit with constant angular momentum its velocity cannot remain constant, since the radius varies at different parts of the orbit. If the velocity of the electron is not constant, then neither will be its mass, according to the Special Theory of Relativity. Hence the energy will depend slightly on the eccentricity of the orbit. This splitting predicted by Sommerfeld's relativistic theory is in very good agreement with experiment.

SELECTION RULES AND THE CORRESPONDENCE PRINCIPLE

When one examines the theoretical and the experimental results for the hydrogen atom, one thing is immediately obvious: not all possible spectral lines are observed. In fact, only transitions between states whose azimuthal quantum numbers differ by unity take place. This is the first example of what is called a *selection rule* that we have come across. To explain the existence of such selection rules, Bohr in 1923 enunciated the *correspondence principle.* This principle is stated in two parts and acts as a bridge between classical and quantum physics:

1. In the limit in which the quantum numbers characterising the state of a system become extremely large, the predictions of the quantum theory for the behaviour of the system must correspond to the predictions of classical physics.
2. Any selection rule holds true over the entire range of the quantum number concerned. Hence any selection rule which is required in order to obtain the required correspondence specified in the first part of the principle in the limit of large quantum numbers also applies in the limit of small quantum numbers.

The first part of the correspondence principle is certainly very attractive. It says, in effect, that we cannot ignore the many successes of classical physics but must impose them as limits on our new quantum theory. It is very plausible that the limit of large quantum numbers should show some correspondence with the classical theory. For, since the unit of quantisation is $h\nu$ if we consider, for example, the energy of an oscillator, then the difference in energy $h\nu$ between the levels of energy $nh\nu$ and $(n+1)h\nu$ for large n is insignificant, i.e. the possible energy levels form a continuum, as they do for a classical oscillator. The second part of the principle is a pure assumption, albeit an attractive one, which explains many of the selection rules observed in the analysis of atomic and molecular spectra. However, its application does lead to some ambiguities.

THE SPECIFIC HEATS OF SOLIDS

In Chapter 1 we introduced a quality called the specific heat* of a system and found that for a mole of a solid the kinetic theory predicted that it should have the value $3Nk \simeq 6$ cal deg^{-1} mole^{-1} in agreement with the observations of Dulong and Petit. However, at low temperatures where the specific heat is observed to go rapidly to zero this agreement is destroyed (see Fig. 25).

The model which leads to the value $3Nk$ for the specific

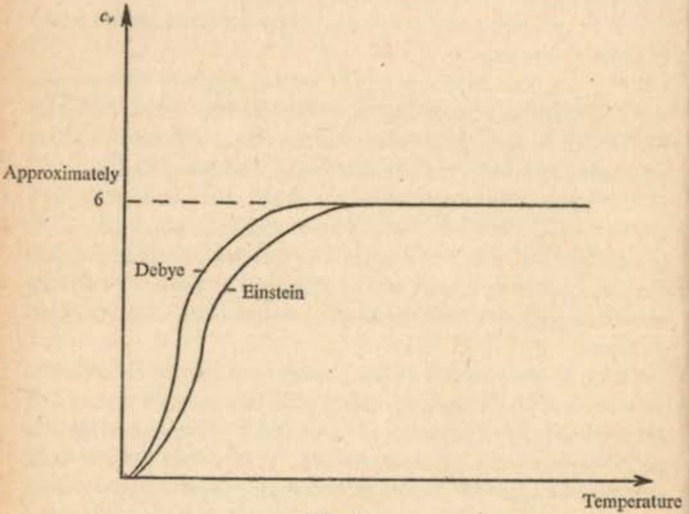


Fig. 25.

---

[*] Throughout, by specific heat, we shall mean the specific heat at constant volume.

heat pictures the solid as consisting of $3N$ linear oscillators, the factor 3 coming from the three-dimensionality of space. Einstein (1906), in attempting to explain the low temperature behaviour of the specific heat, considered the effect of quantising the energy of $3N$ linear oscillators assuming that they all vibrate with the same frequency ν. The result of this analysis is that the average energy of an oscillator is not kT but $h\nu[\exp(h\nu/kt)-1]^{-1}$ (Appendix B) and hence the specific heat c_v becomes

$$c_v = 3Nk\left(\frac{h\nu}{kT}\right)^2 \frac{\exp(h\nu/kT)}{[\exp(h\nu/kT)-1]^2}. \tag{5.10}$$

While (5.10) gives the correct qualitative description of the fall-off in the value of the specific heat at low temperatures, it is not in very good quantitative agreement with experiment. However, a minor modification to Einstein's model made by Debye brought the theory into brilliant agreement with experiment. Debye replaced the condition that all the oscillators should have the same frequency by the condition that the oscillators could have any frequency up to some maximum frequency.

In the limit of high temperatures where kT is very much greater than $h\nu$, both the Einstein and the Debye results reduce to the classical expression $3Nk$.

CRITIQUE OF THE OLD QUANTUM THEORY

The theory which we have outlined in this chapter is now known as the old quantum theory. It has been extremely successful in explaining the phenomena that we have described and which were completely inexplicable in terms of classical physics. However, the theory is altogether unsatisfactory; it is based on a series of seemingly *ad hoc* postulates and, for this reason, if for no other, it appears intellectually unsatisfying. There are a number of other unsatisfactory features of this theory. For instance, the theory only tells us how to treat systems whose co-ordinates are periodic in time; but there are many physically interesting systems which are not periodic. Although the theory tells us the frequency of the quanta emitted or absorbed when a system changes from one quantum state to another, we have no way of calculating the rate at which such transitions take place. We are also left with the baffling evidence of both a corpuscular and a wave nature of light, and with no explanation of this strange duality.

In the years following 1924 these questions received careful scrutiny, and out of the old quantum theory there arose modern quantum mechanics as developed by de Broglie, Schroedinger, Heisenberg, Pauli and Dirac. It is with modern quantum mechanics that we deal in the next chapter.

6. Quantum Mechanics

PARTICLES, WAVES AND THE UNCERTAINTY PRINCIPLE

We have seen in Chapter 2 some very convincing evidence for the belief that light is a wave propagation. Then, in Chapter 5 we have seen that certain phenomena require for their description that we should ascribe a corpuscular nature to light. But we have seen slightly more than this: we have found that there is a definite relationship between the energy, E, of a particle, in the corpuscular description, and the frequency, ν, in the wave description given by

$$E=h\nu. \tag{6.1}$$

De Broglie, in his doctoral thesis of 1924, went even further. He postulated that there was also a definite relationship between the momentum, p, of the particle and the wavelength, λ, of the wave given by

$$\lambda=\frac{h}{p}. \tag{6.2}$$

This can be made plausible by noting that

$$\lambda=\frac{c}{\nu}=\frac{hc}{E} \tag{6.3}$$

and that the energy carried by an electromagnetic wave is related to the associated radiation momentum by $E=pc$. But de Broglie did not stop there; he went on to suggest that the relationships (6.1) and (6.2) were not peculiar to light but were also true for matter.

Davisson and Germer in 1927 set out to test the apparently wild suggestion. The strongest evidence for the wave nature of light is the existence of interference and diffraction patterns. These occur when the dimensions of the diffraction slit, for example, are of the same order, or smaller, in size than the wavelength of the light. If we consider electrons with a kinetic energy of approximately 10 eV, their velocity is much smaller than the velocity of light, and we can calculate the momentum from the classical expression.

$$p = mv = \sqrt{[2m(\tfrac{1}{2}mv^2)]} = \sqrt{(2m \times \text{kinetic energy})}$$
$$\simeq 1{\cdot}7 \times 10^{-19} \text{ g cm s}^{-1},$$

and hence, according to de Broglie, the corresponding wavelength is

$$\lambda = \frac{h}{p} \simeq 4 \times 10^{-8} \text{ cm}.$$

This is an extremely small wavelength compared with that of visible light, approximately 1000 times smaller, so we must look for diffraction slits 1000 times smaller than the usual optical diffraction grating. Nature is extremely kind and provides just such a grating. The spacing between atoms in a solid is of the order of 10^{-8} cm, and hence we can use the aligned rows of atoms in a crystal as a diffraction grating to study the diffraction of electrons or X-rays.

Davisson and Germer carried out such a study and confirmed de Broglie's hypothesis; the diffraction pattern obtained by electrons was of exactly the same form as that obtained by using X-rays.*

We are indeed now faced with a paradox. We all know, at least intuitively, what we mean by a particle or a wave, and surely nothing can truly exhibit both sets of properties. But here we see that if we carry out one class of experiments we appear to observe properties which we associate with particles, e.g. discrete masses, discrete charges, definite energies

and momenta, etc., and on the other hand, if we carry out another class of experiments we observe diffraction and interference patterns typical of a wave motion. If we are to hope to set quantum theory on a firm foundation, we must first understand and resolve this paradox. In order to do this, let us analyse in some detail what we mean by a particle. Above all else, we picture an object localised within a very definite region of space. To define this localised region of space we would, in principle, place some measuring scale alongside the particle and read off the exact measurement of its front and back edges relative to some reference point. Then, in relation to this reference point, we would say that we knew the position of the particle exactly. Suppose we now try to do this with an extremely small object. To detect the particle we must look at it, i.e. light is scattered off the object and into our eyes, and hence we see the particle. As the object gets smaller and smaller, we shall require a shorter and shorter wavelength in order to see it; many optical microscopes are illuminated by blue light (short wavelength) to increase their resolution. Now, as the wavelength becomes shorter and shorter, according to the de Broglie postulates the momentum of the light corpuscles increases. Hence we can increase the accuracy of our knowledge of the position of a particle at a given time as much as we may wish, but only at a price. The price we must pay is a loss of information about the energy and momentum of the particle, because as the energy and momentum of the light corpuscle increases so does the violence of the scattering process. It must be stressed that this is not simply a difficulty in practice but a limit in principle set by nature on our knowledge of the universe. Having the senses which we possess and which are our only windows on the universe, there is no way in which we can overcome this balance of information. We can, of course, extend our senses by using electron microscopes, etc., but on analysis we always find the same limit on the maximum knowledge which we can in principle obtain about a system.

Naturally, these considerations only become important when we study objects of subatomic dimensions. One can hardly imagine a torch beam significantly altering the momentum and energy of a railway engine. In other words, just as the theory of relativity only became important in systems moving at velocities near the velocity of light, the problem of the uncertainty of the momentum of an object only becomes important in objects of the dimensions of the wavelength of light, i.e. at the extreme limits of our senses.

---

* We have used here the diffraction of electrons through a crystal lattice, whose spacing is assumed known, to observe the wave motion and calculate an associated wavelength. Obviously this process can be inverted; knowing the wavelength of, say, X-rays, by studying the diffraction pattern we can obtain information about the spacing between the atoms of the crystal. We shall not have space in this small volume to discuss the subject of X-ray crystallography, but would just mention that it is from considerations of this type that the value of Avogadro's number used in Chapter 2 could be estimated.

Now, in fact, we always have some information about a particle we may wish to study. Even a simple piece of information such as the fact that the electron is somewhere in the laboratory immediately places some restriction on our knowledge of its momentum. Admittedly, the restriction in this case would be very weak, but, none the less, it is a real restriction. If this is so, then an electron can hardly be described as a particle, because this immediately implies that it can be localised in some specific region of space and have ascribed to it a definite energy and momentum.

What is required in order to describe an electron, or any other subatomic particle, is something which has built into it this 'in principle' impossibility of measuring its position in space and time and its momentum and energy simultaneously to arbitrary accuracy. Not only this, we should also like a description which automatically accounted for the observed wave-particle duality. Such a problem has been familiar for a long time in mathematics, and its solution forms that branch of mathematics known as Fourier analysis, where it is shown that given any arbitrary shaped curve it can be represented by a linear combination of wave curves of varying amplitudes and wavelengths. However, if we wish to specify our arbitrary shaped curve exactly, then in general we shall require an infinite number of wave curves, the number of wave curves required increases as we increase the accuracy with which we wish to represent our arbitrary curve. If with each wave curve we associate the de Broglie relations (6.1) and (6.2), then we see that we have exactly duplicated the problem which we face when we attempt to measure the position and momentum of a particle, and thus we can use this type of analysis to describe the behaviour of subatomic particles. We shall represent our particles by groups of waves and call such an object a *wave packet*. If we try to localise our wave packet in a very definite region of space and time, it will require many waves to represent it, each with a different wavelength (momentum) and frequency (energy), i.e. our knowledge of the momentum and the energy of the wave packet will become blurred. On the other hand, if we measure its energy and momentum very accurately; that is, we can represent it through the de Broglie relations by a single wave, we shall have to forfeit our knowledge about its position in space and time, since a single wave is not localised.

Fourier analysis tells us slightly more than this, it also sets a lower limit on the number of waves, i.e. range of momenta (wavelengths) Δp required to specify a region with an accuracy Δx, it is given by

$$\Delta x \Delta p \geqslant h/2\pi. \tag{6.4}$$

There is a corresponding relationship in the range of energies (frequencies) ΔE required to specify an event within a time Δt

$$\Delta E \,.\, \Delta t \geqslant h/2\pi. \tag{6.5}$$

The equations (6.4) and (6.5) are expressions of the Heisenberg Uncertainty Principle, and set limits on the accuracy with which either the momentum and the position or the energy and the time of a wave packet can be known simultaneously. We can see now why classical physics was only found inadequate at the microscopic limit of the human senses, for h is an extremely small quantity (approximately $6{\cdot}6 \times 10^{-27}$erg s^{-1}) and hence the uncertainty principle is only important where the quantities which we are measuring are extremely small. In fact, if h were zero, classical physics would have given a correct description of the subatomic world.

By talking about wave packets instead of particles, we are provided with the answer to another difficulty which we discussed in Chapter 2. We indicated that one of the crucial points in deciding between Newton's corpuscular theory and the wave theory of light was their predictions concerning the refraction of light. You will remember that Newton's theory required that the corpuscles moved faster in a denser medium, while the wave theory required that light travel slower in denser media. From Fig. 26 we see that Newton and the wave theorists were, in fact, talking about different velocities, and we must now distinguish between the group or particle velocity and the phase or wave velocity of the individual waves making up the wave packet. The particle velocity is the group velocity in the case of an electron, for instance. The group velocity is also the velocity at which the energy and momentum are transmitted in the case of an electromagnetic wave. It can be shown that the group or particle velocity and the wave velocity are inversely proportional, and hence we have an explanation for the apparently contradictory predictions of the two classical theories.

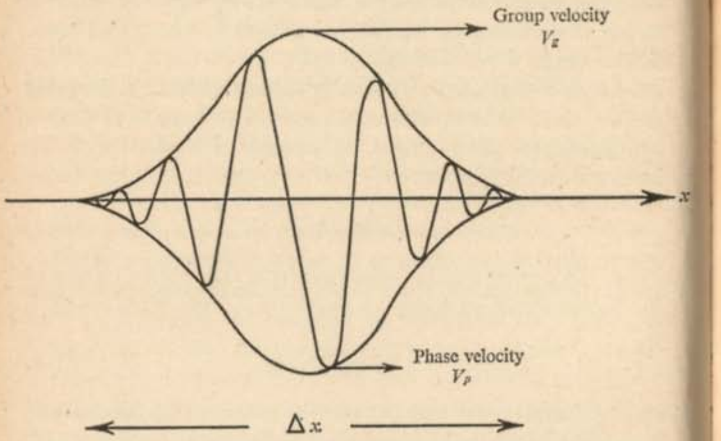
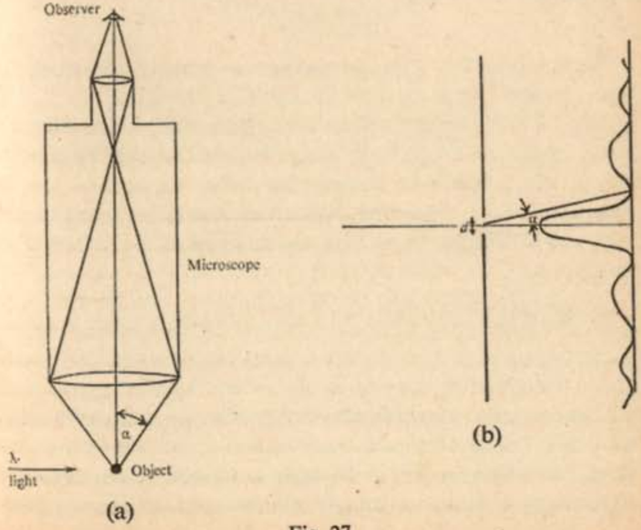


Fig. 26.

This is all very well. So perhaps we cannot simultaneously measure the momentum and position of a wave packet, but we do know that if we have a single electron, or photon, and pass it through a diffraction grating, we are not going to obtain a complete diffraction pattern; all we shall obtain is a single point on the screen where the particle strikes it. However, if we take a beam of identical particles and shine them through a diffraction grating, then we will obtain the appropriate diffraction pattern, and this gives us a clue to the interpretation of the meaning of the wave packet. The diffraction pattern is nothing more or less than the statistical distribution of the particles on the screen. Although the particles are all identical they each carry an uncertainty in their exact position and momentum, and this manifests itself in a statistical scatter of the positions at which the particles strike the screen; it is just this scatter which we call the diffraction pattern. Hence the wave packet can be said to represent the probable properties of the particle, i.e. the spread of the wave packet sets upper limits on the probable position of the particle, the range of waves required to represent the wave packet give the probable values of the energy and the momentum of the particle.

In a number of simple cases we can show the relationship between the optical effects (statistical features) and the Heisenberg Uncertainty Principle. In Fig. 27(*a*) we schematically



(a)

(b)

Fig. 27.

represent an object being viewed through a microscope. From classical optics the resolving power of the microscope, i.e. essentially the accuracy with which its position can be established is $\lambda/\sin\alpha$, where λ is the wavelength of the light illuminating the object. The total momentum of the light is $p=h/\lambda$, and hence the resultant uncertainty in the momentum of the object as a result of the light striking it is $2p\sin\alpha$.

$$\Delta x \simeq \lambda/\sin\alpha$$

$$\Delta p \simeq 2p\sin\alpha = (2h/\lambda)\sin\alpha$$

$$\Delta x \Delta p \simeq 2h.$$

In Fig. 27(*b*) we represent the diffraction pattern produced by a beam of light passing through a single slit of width d. Since the photons pass through the slit, we know their position within a distance d. Since the particle could strike the screen anywhere within any of the maxima, the uncertainty in its momentum is at least as great as that required to produce the width of the first maxima, i.e. $\Delta p \geqslant 2n\sin\alpha/\lambda$. Thus $\Delta x\Delta p \geqslant (2d\sin\alpha/\lambda)h$. But again, by classical optics, $2d\sin\alpha=\lambda$ and hence

$$\Delta x\Delta p \geqslant h.$$

We are now left with the problem of how to construct the wave packet for a particle in some given situation. It will suffice us here simply to state that two prescriptions were set forth by Schroedinger and Heisenberg in the years 1925–26. Since then it has been shown that two prescriptions are, in fact, equivalent. However, the mathematical sophistication required to follow these procedures is beyond the scope of this book.

SPIN, THE EXCLUSION PRINCIPLE AND THE STRUCTURE OF ATOMS

First, let us return to Bohr's theory of the hydrogen atom. Here the electron moves in states of definite energy and momentum, i.e. they must be describable in terms of a wave of definite frequency and wavelength. If at any space-time point the wave packet is to have a unique value, then the orbit must contain an integer number of wavelengths (Fig. 28), i.e. the circumference of the classical orbit is $2\pi r=n\lambda$, but the momentum of the particle is given by $p=h/\lambda$ and hence $2\pi r=nh/p$. Now, the angular momentum of a particle in a circular orbit is simply $L=pr$, and thus

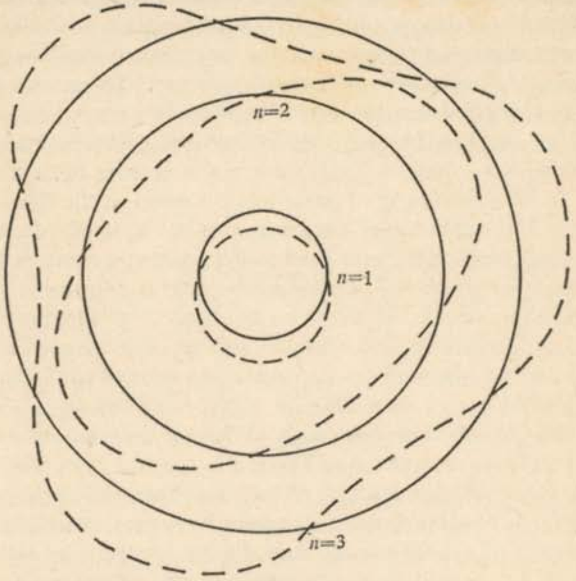


Fig. 28.

$$L=\frac{nh}{2\pi}$$

which is simply Bohr's quantisation rule.

Now let us consider the structure of a many-electron atom. Our first attempt to construct such an atom might be to put all the electrons in the lowest Bohr orbit; this, in fact, does not happen. Classically we might argue that this is because the electrons, all having the same charge, repel one another, and thus we might expect some extended distribution of the electrons over the Bohr orbits providing conditions of electromagnetic stability. We would further argue that these Bohr orbits would not be the same as those of the one-electron atom, but would be determined not only by the field of the nucleus but also by the field of the electron distribution itself leading to conditions of stability. Such a problem is called a self-consistency problem, and the calculation was initially attempted classically by Thomas and Fermi. The results are not in very good agreement with experiment.

If we examine spectral lines in a strong magnetic field, it is found that what are single lines in the absence of the field are, in fact, multiplets of lines; this is known as the Zeeman effect. Some splitting of the lines is to be expected, because although each orbit has a fixed total angular momentum the direction in which the angular momentum is pointing is not fixed. Now, since an electron moving in an orbit behaves like a small magnet (p. 23) and the energy of a magnet in a magnetic field depends on its orientation relative to the field, we should expect that different orbits with the same total angular momentum but with different orientations will have different energies and hence the spectral lines will be split. However, this splitting of the parent spectral line is not completely random, but is the same in every case. A transition to a state of orbital angular momentum L, where $L = lh/2\pi$, $l = 0, 1, 2, \ldots$, will always be split into $(2l+1)$ spectral lines when placed in a magnetic field. This can be explained as follows; not only is the total orbital angular momentum quantised but, in addition, the orientations of the orbits are also quantised. This can be put in another way: suppose L_z is the component of the total angular momentum along some fixed direction, for instance, along the direction of the magnetic field, then L_z is also quantised, i.e.

$$L_z = m_l h/2\pi \quad m_l = l, (l-1), (l-2), \ldots, -(l-1), -l, \tag{6.6}$$

and we see then, that there are, indeed $(2l+1)$ possible values of m_l.

In 1922 Stern and Gerlach set out to measure the possible values of m_l by sending a beam of atoms through a very strong, non-uniform magnetic field. Classically the beam should have been continuously spread out by the magnetic field corresponding to the equal probability of all possible orientations of the electronic orbits. Quantum mechanically we should expect the beam of atoms to be split into $(2l+1)$ parts corresponding to the quantisation of the orientations of L. However, when a beam of silver atoms which have zero orbital angular momentum passed through the field, instead of remaining as a single beam it was split into two parts (see Fig. 29). Uhlenbeck and Goudsmit both qualitatively and quantitatively explained this strange splitting by suggesting that the electrons themselves spin about their own axis as they rotate around the nucleus in exactly the same way as the Earth spins about itself, giving night and day, while rotating around the sun giving rise to years. The electron spin is, however, not arbitrary but is itself quantised, having only one value $S = \frac{1}{2}h/2\pi$. The orientation of the spin is also quantised, having $(2 \times \frac{1}{2} + 1)$, or simply 2, values $S_z = \pm\frac{1}{2}h/2\pi$, and hence giving rise to the observed splitting into two parts.

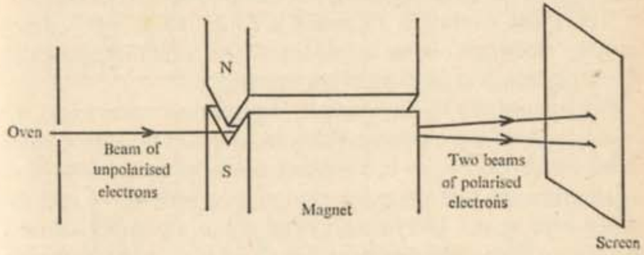

N
Oven
Beam of unpolarised electrons
S
Magnet
Two beams of polarised electrons
Screen

Fig. 29.

We see now that whereas in Bohr's simple model of the atom the orbits of the electrons were labelled by a single quantum number n, the principal quantum number, we really require also for the complete specification for an electron state the orbital angular quantum number l, a component of the orbital angular momentum m_l, a spin quantum number $S(=\frac{1}{2})$ and a component of the spin $m_s(=\pm\frac{1}{2})$. It is then an empirical fact that no two electrons in the same system can ever have the same complete set of quantum numbers; this rule is known as the Pauli Exclusion Principle. The exclusion principle immediately suggests how nature constructs many-electron atoms. The electrons fill the Bohr orbits one by one, and only as many are allowed into each orbit as the exclusion principle will admit. Thus each Bohr orbit of angular momentum quantum number l will contain a maximum of $2(2l+1)$ electrons, the factor 2 comes from the two possible orientations of the spin and the factor $(2l+1)$ corresponds to the possible orientations of the orbital angular momentum, then any additional electrons must start to fill the next orbit.

The Pauli Exclusion Principle is found to be obeyed not only by electrons but by all particles with a half integer value for their intrinsic spin quantum number.

In terms of the Pauli principle, we can now understand the regularities of the Periodic Table (Fig. 30). The lowest Bohr orbit has $n=1$, $l=m=0$, $s=\frac{1}{2}$ and hence is completely filled when there are two electrons present, one with $m_s=\frac{1}{2}$ and the other with $m_s=-\frac{1}{2}$. The next orbit has $n=2$ with $l=0$ or 1 and hence is completely filled when it contains eight electrons, and so on. Thus the vertical columns in the Periodic Table contain elements with similarities in their electronic structure and hence their chemical properties, e.g. below hydrogen we have the alkali metals lithium, sodium, potassium, etc., all of which have strong chemical reactions and contain a single electron outside a series of completely filled orbits. Under helium we have argon, krypton and the rest of the inert gases. These all have closed shells, i.e. all occupied orbits are completely filled and they have very weak chemical reactions and hence the generic term inert. Thus we see that the Periodic Table is nothing but a classification of the elements according to their electronic structure.

The inert gases will be electrically neutral and, consisting solely of closed shells, they will be spherically symmetric; hence there will be no electric or magnetic forces between such atoms, and consequently they will only interact very weakly. The alkali metals, on the other hand, have a single electron outside a closed shell which can easily be removed, or, as we say, the atom can be *ionised*. The resultant ion will then have a net positive charge (having lost a negative electron)

THE PERIODIC TABLE OF THE ELEMENTS

Fig. 30. The number of electrons is given in the upper left corner. Some sample atomic weights are given in the lower right corner.

n=1	1 H 1.008																	2 He 4.003
n=2	3 Li 6.94	4 Be 9.02											5 B 10.8	6 C 12.01	7 N 14.01	8 O 16	9 F 19	10 Ne 202
n=3	11 Na	12 Mg											13 Al	14 Si	15 P	16 S	17 Cl 35.5	18 A
n=4	19 K	20 Ca	21 Sc	22 Ti	23 V	24 Cr	25 Mn	26 Fe	27 Co	28 Ni	29 Cu	30 Zn	31 Ga	32 Ge	33 As	34 Se	35 Br	36 Kr
n=5	37 Rb	38 Sr	39 Y	40 Zr	41 Nb	42 Mo	43 Tc	44 Ru	45 Rh	46 Pd	47 Ag	48 Cd	49 In	50 Sn	51 Sb	52 Te	53 I	54 Xe
n=6	55 Cs	56 Ba	57 ↑ La	72 Hf	73 Ta	74 W	75 Re	76 Os	77 Ir	78 Pt	79 Au	80 Hg	81 Tl	82 Pb	83 Bi	84 Po	85 At	86 Kn 222
n=7	87 Fr 223	88 Ra 226	89 ↑ Ac 227															
	$l=0$		$l=2$										$l=1$					

58 Ce	59 Pr	60 Nd	61 Pm	62 Sm	63 Eu	64 Gd	65 Tb	66 Dy	67 Ho	68 Er	69 Tm	70 Yb	71 Lu
90 Th	91 Pa	92 U	93 Np	94 Pu	95 Am	96 Cm	97 Bk	98 Cf	99 Es	100 Fm	101 Md	102 No	

$l=3$

and hence will interact strongly through its electric field especially with elements like chlorine which require only a single electron to complete a closed shell, and hence borrow the sodium electron and acquire as a result a net negative charge. The positive sodium ion then binds strongly to the negative chlorine ion through the ordinary electrostatic attraction of unlike charges to form sodium chloride (common salt). In solution, the binding between the ions is not as strong as it is in the solid state, and the compound molecules can then be made to dissociate by applying an electric field, hence the phenomena of electrolysis. Sometimes, as in the case of magnesium oxide, two electrons are exchanged in order to produce the closed shell configuration, and in this case the binding between the ions is, as we might expect, four times as strong. On other occasions it is advantageous for an atom with more than one electron missing from a closed shell configuration to borrow electrons from a number of atoms with electrons in excess of a closed shell structure, e.g. an oxygen atom combines with two hydrogen atoms to form water.

The electron outside a closed shell, or the holes caused by electrons missing from a closed shell, are called valency electrons or holes, and in terms of these much of modern chemistry finds an explanation.

Let us return to a consideration of the exclusion principle. Particles which obey this principle are called *fermions*, after the Italian physicist Enrico Fermi. These particles appear to be the building blocks out of which the universe is constructed, e.g. electrons, protons and others which we shall meet briefly in the last chapter. Particles which do not obey the exclusion principle are called *bosons*, after the Indian scientist Bose. The bosons have integer multiples of $h/2\pi$ for the value of their spin. (For the statistics of fermions and bosons see Appendix B.) The only bosons that we have met so far are photons,* the particles of light; they have spin $1 \times h/2\pi$. Having unit spin means that there are three possible orientations of the spin $m_s = 1, 0, -1$. $m_s = \pm 1$ corresponds to the existence of right and left circularly polarised light, and phonons with $m_s = 0$ do not in fact arise in nature. The elementary bosons appear to be associated with the fields that bind the fermions together, e.g. photons are associated with the electromagnetic field through which charged particles interact. In the last chapter we shall meet the *pions*, which appear to be the analogue of the photons and are associated with the nuclear field which holds particles together to form atomic nuclei.

The exclusion principle appears to be nature's way of carrying out her book-keeping; nature likes to know how many fermions there are in the universe. It is rather like an extremely short-sighted person peering into a room which contains two identical twins, how could he ever know that there are, in fact, two twins? If he waits long enough, one of the twins will eventually come within his range of vision and then pass out of it again; some time later he will see another twin. But how does he know it is not the twin he has already seen? If the twins are truly identical, the observer can never tell whether he is seeing the same one over and over again or whether there are many of them. In the same way, if nature allowed more than one fermion to have exactly the same quantum numbers, she could never tell how many fermions there are. However, if the twins are in different rooms (the fermions in different quantum states), the observer can go first to one room and see the first twin and then to the second room and hence see the second twin, and in this way decide that there are two of them.

The last point that I wish to make in this discussion of quantum mechanics is that since the wave packet gives us a measure of the probability of finding the particle at a given point in space at a given time and with a given momentum and energy, we can calculate the probability of a transition from an initial state to some final state and hence predict the rates of various electronic transitions in atoms, etc. By and large, these are found to be in excellent agreement with experiment.

---

* We have, of course, also the α-particle which has zero spin, but at the risk of being out of step with popular modern usage we shall not consider the α-particle as being as elementary as the photon and other bosons to be discussed later.

7. Pot-pourri

ELEMENTARY PARTICLES

Prout in 1815 had suggested that all elements were, in fact, built out of hydrogen atoms. A quick glance at the Periodic Table (Fig. 30) will suffice to convince that this cannot be so, since, if it were, all elements would have atomic weights which would be integers on a scale in which the atomic weight of hydrogen was unity, e.g. chlorine with an atomic weight of 35·5 contradicts this suggestion. Another observation which is in conflict with Prout's hypothesis is that while the atomic weight of helium is four times the atomic weight of hydrogen, the nuclear and electronic charges of helium are only twice what they are in the case of hydrogen. A possible explanation of these phenomena is that atomic nuclei, besides containing protons (hydrogen nuclei), also contain other particles which have approximately the same mass as protons but are electrically neutral. Such particles, called *neutrons*, were suggested by Rutherford, and empirical evidence for their existence was put forward by Chadwick in 1932 after studying some experiments by Curie and Joliot. Since the neutrons are electrically neutral, they cannot affect the chemical behaviour of an atom. Hence we could construct atoms with the same number of electrons and photons but with different numbers of neutrons, and such atoms, while having different atomic weights, would be chemically indistinguishable; such atoms are called isotopes of the same element. By mixing different isotopes into a chemical sample it is obvious that we can obtain average atomic weights which are not integer multiples of that of the hydrogen atom (e.g. the isotope of chlorine with atomic weight 35 occurs approximately 80 per cent of the time in nature, while another isotope with atomic weight 37 accounts

for the remaining 20 per cent of the time, the resultant average atomic weight is 35·5, as indicated in Fig. 30).

Carbon is taken as the reference atomic weight, because it possesses one isotope containing six protons and six neutrons which is very much more abundant than any other. On the scale with carbon as 12, hydrogen has an atomic weight of 1·008 corresponding to the existence in nature of small but measurable quantities of the hydrogen isotopes deuterium (one proton plus one neutron) and tritium (one proton plus two neutrons).

The neutron, like the proton and the electron, turns out to be a fermion with intrinsic spin $\frac{1}{2}h/2\pi$.

The electron and the proton were for many years considered to be the only truly fundamental particles, but now the neutron appears on the scene as a new contender for this title. However, there is one great difference which distinguishes the neutron from the electron and the proton; whereas the electron and proton are stable, i.e. if left to themselves they will always remain as an electron or a proton, the neutron will decay into a number of new particles in, on average, approximately 12 minutes. The most obvious products of this decay are an electron and a proton, and the immediate reaction to this is to describe the neutron as a bound state of an electron and a proton and return to the situation where there are only two elementary particles. However, there must be another product of the neutron decay, because the electron and the proton both being fermions with spin $\frac{1}{2}h/2\pi$ could only combine to form a particle with spin of either $0h/2\pi$ or $1h/2\pi$, i.e. a boson. The other product of the decay must indeed be a strange particle; since the electron and the proton satisfy the charge and mass requirements of the neutron, it can have very little mass and must be electrically neutral. On the other hand, it must possess an intrinsic spin of $\frac{1}{2}h/2\pi$, i.e. be a fermion. This particle is called a *neutrino*; it turns out to be stable and its rest mass appears to be identically zero.

At this stage we would appear to have four fundamental particles, if we neglect the neutron. There is one boson, the photon, and three fermions: the neutrino, the electron and the proton. As if this was not enough, in 1938 the Japanese physicist Yukawa postulated the existence of a new particle, to be called the meson. This particle was to be a boson and was to play the same role in connection with the field of force which holds the nucleus together that the photon plays in the electromagnetic field which causes the interaction between charged particles. Yukawa, from considerations of the range and strength of nuclear forces, predicted that the meson would have a mass of approximately 273 electron masses. In 1939 a new particle was discovered as a product of cosmic rays (radiation, sometimes of extremely high energies, falling on to the Earth from outer space). It was immediately assumed that this new particle was Yukawa's meson and it was named the μ-meson, or simply *muon*. However, a closer study revealed that this new particle was a fermion with a mass of approximately 240 electron masses and a negative charge. In many respects this muon behaved like an overgrown electron. It was an unstable particle decaying in a time of the order of 2×10^{-6} seconds into an electron and two neutrinos. It most certainly did not interact strongly with either the neutron or the proton, as it should have done were it the Yukawa meson. And so the search began anew for the Yukawa particle now known as the π-meson, or simply the pion; eventually it was discovered by Powell in 1945 at Bristol. But the proliferation of particles does not end here, for instead of there simply being one pion there turned out to be three distinct types: two of them carried electric charges identical to those found on the electron and the proton, while there maining one was electrically neutral.

Since the Second World War, especially with the advent of the giant particle accelerators, the number of known particles has multiplied many times. Many new heavy fermions and heavy bosons have been discovered, and even old friends have revealed unsuspected depths. For example, it appears that there is not one neutrino but two; and in case you worry how there can possibly be two, electrically neutral massless objects spinning about themselves which can be distinct, the property which distinguishes them is that one set will only

interact with electrons, while the other prefers the company of muons.

The nightmare is not yet over, because if we look back to equation (3.34) we see that there are, in fact, two solutions

$$E=\pm\sqrt{(c^2p^2+m_0^2c^4)}. \qquad (7.1)$$

Initially the negative solution was simply ignored as being physically meaningless; after all, no one has ever observed a free particle in a negative energy state. Dirac suggested that this was because all the negative energy states were filled and the Pauli Exclusion Principle would then prohibit any other fermions from entering these states. In other words, the universe is set against a back-drop of a sea of full negative energy states, and we can only observe events which lie above this sea, i.e. particles in positive energy states. However, if we had enough energy we should be able to pluck a particle out of this sea and raise it to an observable positive energy state (Fig. 31). Such a particle would leave behind a hole in the negative sea. This hole, or *antiparticle*, as it is now called, will behave in all respects like the particle which used to fill it, except that all its characteristic properties will be reversed. In 1932 Anderson discovered the first antiparticles while studying cosmic ray showers; this was the *positron*, or positive electron.

The energy required to produce an antiparticle is $2m_0c^2$ (or approximately 1 MeV in the case of electrons). Such enormous energies are now available, and in fact the construction of a machine capable of generating 300 000 MeV is at present being contemplated, and many heavier antiparticles have been produced. When a particle and an antiparticle collide, both particles vanish, leaving behind only a vast amount of radiant energy.

It is no longer fashionable to talk about negative energy states but rather to use phrases like 'particles going backwards in time' or indeed to give up all attempts at providing a descriptive picture of events and simply talk about positive and negative frequency solutions of the relativistic wave equation.

For many years the vast number of fundamental particles provided a bewildering spectacle. Most of the particles are extremely unstable and decay into other particles, which in turn also decay – the whole process of continual interchange going on continuously until the whole picture of what is truly fundamental and what is merely transitory appears to lose all meaning.

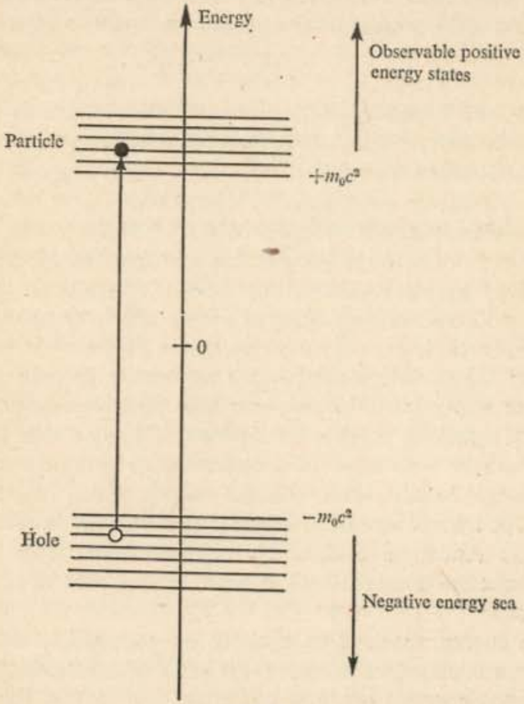


Fig. 31.

Within the last few years a new pattern has slowly emerged from this apparent chaos, at least for the baryons, i.e. the neutron and proton and heavier fermions, together with the heavy bosons which take part in the interactions between these fermions. Following a suggestion by the Japanese physicist Ohnuki, Ne'eman, Gell-Mann and Salam have presented us with a classification scheme for the baryons. Using this classification scheme, it has been possible to predict the existence of, and properties of, several new particles which have subsequently been observed in high energy experiments. It is rather like having the Periodic Table of Elements without knowing anything about atomic structure. Before we can hope to answer any of the questions as to why the classification scheme should be as it is, we shall require to know a lot more about the actual structure of the elementary particles. As always, the most difficult questions remain unanswered. Why do the particles have the discrete masses that they do have? Why do they spin as they do, and why is there a unit of spin? Why are they electrically charged, and why is there a unit of electric charge? And, finally, what, if any, is the relationship between the four basic forces in the universe: gravity, weak (the force through which the neutrino and the electron interact), electromagnetic and nuclear?

FISSION AND FUSION

In the previous section we mentioned nuclear forces frequently. That there is a need for such forces is obvious, because if they did not exist the electrostatic repulsive forces between the protons would blow every nucleus in the universe to pieces. From the simple observations which we have already made, we can tell a considerable amount about the force between nucleons (neutrons or protons). Firstly, it has a much shorter range than the electromagnetic force, because the α-particles in Rutherford's scattering experiment were repelled by the nuclei. However, the nuclear force must become much stronger, and attractive, at shorter distances in order to combat the electrostatic repulsion. Finally, at very short distances the nuclear force must become repulsive, or else all nuclei would simply collapse as the nucleons came closer and closer together. Experiments suggest typical values of $1{\cdot}5 \times 10^{-13}$ cm for the range of the attractive part of the nuclear force, and $0{\cdot}5 \times 10^{-13}$ cm for the range of the repulsive core. The force between two neutrons or a neutron and a proton appears to be identical to the force between two protons, apart from the ordinary electrostatic repulsion. Thus, adding neutrons to a nucleus simply increases the overall binding of the nucleus, while adding a proton may help the binding in the vicinity of the proton but will act as a repulsion to another proton on the far side of the nucleus which is out of range of the first proton's nuclear attraction. From considerations of this type, one can calculate the equilibrium number of neutrons and protons in any nucleus. The result is that for many light nuclei, from helium with an α-particle nucleus, i.e. two protons and two neutrons, to calcium with 20 protons and 20 neutrons, the equilibrium configuration contains as many protons as neutrons. However, as the mass increases so does the effect of the electrostatic repulsions, and hence in heavier nuclei an excess of neutrons over protons is required for equilibrium, e.g. uranium has 92 protons but approximately 146 neutrons, depending on the isotope considered. If a heavy nucleus is unstable, there are three methods by which it can try to establish stability. Firstly, if there are too many neutrons and not enough protons it can adjust the balance by converting one or more of its neutrons into protons and emitting the resultant electrons; this is known as β-decay and is the source of the β-rays discussed in Chapter 4. Secondly, it can increase the ratio of the number of neutrons to protons by emitting an α-particle. Finally, in some cases., e.g. uranium, the nucleus may *fission*, i.e. split into two parts. Both daughter nuclei, being much smaller than the parent, will require fewer neutron for their stable configurations, and hence in addition to the two daughter nuclei there will be a number of free neutrons created. Furthermore, both the daughter nuclei will be much more tightly bound (i.e. stable) than their parent; that is, there will be a great excess of potential binding energy, and this is given out in the form of kinetic energy for the neutrons. In a small piece of fissionable material the product neutrons will have passed out of the material before they encounter another nucleus. However, if the piece of fissionable material is large enough, the neutrons produced in any one fission are likely to collide with another nucleus before escaping from

the material. As a result of this collision, the second nucleus will fission and hence a chain reaction can be built up extremely rapidly. The limiting size of a piece of fissionable material below which a chain reaction will not take place is known as the critical mass.

In Chapter 2 we saw that heat energy was just a measure of the internal kinetic energy of the particles in the system. Hence in a chain reaction in which fission products are produced with high kinetic energies we generate a great deal of heat very quickly.

In an atomic bomb this heat generation is uncontrolled. Two subcritical masses of uranium are separated by a neutron-absorbing shield of lead. If the lead is removed and the pieces of uranium are forced together to form a total mass in excess of the critical mass, a chain reaction will set in and a vast quantity of heat will be generated, together with the emission of a large neutron flux characteristic of an atomic explosion. A more useful application of the fission process is to build a structure of subcritical fissionable masses separated by non-fissioning, neutron-slowing shields (e.g. carbon) with a geometry such that the rate at which heat is generated remains constant and at a level at which it can be used to drive a turbine. Such a structure is called an atomic reactor, and in one form or another it is the basis of modern nuclear power stations.

Another way of converting potential binding energy into kinetic energy, and hence heat, is by *fusion*. The fusion process takes place at the opposite end of the mass spectrum from fission. Since, for instance, the α-particle is such a stable structure, if four hydrogen atoms or two deuterium atoms can be persuaded to fuse into an α-particle, we shall gain a great deal of potential binding energy which will be in the form of radiant energy and kinetic energy. The number of particles taking part in the fusion process is extremely small, and hence the share of kinetic energy per particle will be much greater than in the fission process. Thus much higher temperatures are reached in a fusion explosion (H-bomb) than in a fission explosion.

Chain reactions based on the fusion process are generally believed to be the source of solar and stellar energies.

To date, no satisfactory method of producing a controlled fusion plant has been announced.

GENERAL RELATIVITY

Throughout this book we have concentrated on the microscopic world; let us now turn briefly to a discussion of the very large.

In his special theory of relativity we saw that Einstein had presented us with a host of new ideas concerning the structure of space and time. However, the theory only applied to inertial systems, i.e. non-accelerated frames of reference. In 1915, in his general theory of relativity, Einstein extended these ideas to systems accelerating relative to one another. This theory, unlike the Special Theory of Relativity, has not acquired a vast amount of experimental evidence in its favour in the 50 odd years since its initial introduction, nor has it yet interacted strongly with any branch of the main stream of modern physics. However, in recent years there has been a growing body of opinion which has suggested that in a detailed study of gravitational theory we may find the answer to some of the perplexing questions confronting us in elementary particle physics.

Einstein considered the following gedanken experiment. A scientist is in the cabin of an elevator which is falling freely under gravity. There are no windows in the cabin, so the scientist cannot tell that he is moving by reference to any external object. Then in this cabin the scientist can carry out any experiment that he may wish, and he will never be able to tell whether he is falling freely in a gravitational field or whether there is simply no external gravitational field at all. In other words, someone in a frame of reference external to the cabin would say there was a gravitational force, while someone in the frame of reference of the cabin would say there was no gravitational force. There is perhaps a better known example of this type of argument: if one is in a frame of reference which is rotating relative to some object, the fictitious centripetal and centrifugal forces must be invented

in order to describe its motion in accordance with Newtonian mechanics. We have said that the centripetal and centrifugal forces are fictitious; they are in the sense that if the frame of reference is chosen carefully they will vanish. Having considered his scientist in the elevator cabin, Einstein came to the same conclusion about gravitational forces; in the correct frame of reference they would vanish. Then the problem of the structure of the universe would not be one of dynamics but simply one of geometry. The correct geometry for this problem, unfortunately, is not Euclidean but Riemannian; that is to say, the space is curved. We should remember that not only is it curved but that it is a four-dimensional space-time continuum. Thus the planets do not move round the Sun in elliptic orbits under the force of gravity, but rather because the space around the Sun is curved and the planets are following geodesics in this curved space.*

Einstein found that with the aid of tensor calculus he could write the laws of physics in a general form valid for all reference frames. Unfortunately, the mathematics do not lead to a unique expression for the curvature of space, i.e. a unique theory of gravity. From amongst the many possible models Einstein selected the one which appeared to him to be mathematically the simplest, and proceeded to study the predictions of this model. In particular, there are three very famous predictions which appear as a novel feature of the new theory. They are:

1. Since space is curved in the proximity of a massive body, light rays passing close to such a body should appear bent. This is, in fact, observed, and stars seen adjacent to the Sun during an eclipse appear to have an angular displacement of 1·75 seconds of arc relative to their position when their rays do not pass close to the Sun.

* A geodesic is to a curved-space geometry what a straight line is to a Euclidean geometry. An airliner does not travel from London to New York in a straight line; such a suicidal path would take it deep beneath the Atlantic Ocean. Instead, the airliner travels along the shortest route between London and New York which is compatible with the surface of the Earth being curved, i.e. along a great circle which is the geodesic for a sphere.

2. Physical processes taking place where the radius of curvature of space is very long (i.e. where space curves very slowly, or, in other words, where there is a high gravitational potential) take place more quickly than at a point with a shorter radius of curvature. For example, atomic vibrations in a massive star would appear to be slower than the same vibrations on Earth and, consequently, line spectra emitted by such stars should appear to be shifted towards the red end of the spectrum.

3. Kepler had demonstrated that a body moving in inverse square law of force (Newtonian gravity) should, if it was a periodic orbit, move in an ellipse. Now, this is clearly only a first approximation in the case of planetary motion around the Sun. Firstly, the Sun is not infinitely heavy, and so rotations will not be around the centre of the Sun but around the centre of mass of the Solar System. Secondly, in considering any particular planet we should include in our considerations not only the effect of the Sun but also of the other planets. The net effect of these perturbations is to produce a rotation of the perihelion of an elliptic orbit; however, the predicted rate of this rotation as given by a classical calculation and by Einstein's general relativity do not agree. Einstein's theory predicts a slightly higher rate of rotation; in the case of the planet Mercury an additional 43 seconds of arc per century, or an extra complete revolution every three million years. Observations support Einstein's theory.

There are several rival theories, many of them minor variations of Einstein's theory, which make similar predictions. The present astronomical observations of very distant objects in our universe should soon allow us to eliminate many of the proposed theories. However, it appears unlikely that empirical observation will present us with a unique theory in the near future, because the differences in the predictions of many of these theories are so fine that it would appear improbable that the accuracy of experimental measurements could be improved sufficiently to decide between them.

A more general point should now be made. In this section we have talked only about laws of gravity, but obviously it is a very attractive hope that some day we might be able to incorporate all laws of force within a single theory. Work on such a *unified field theory* has been attempted ever since Einstein's General Theory of Relativity was put forward. It is only natural that the first attempts should be to produce a single theory accounting for both gravity and electromagnetism; after all, their classical laws of force are identical in form. Over the years, little success has been recorded in this difficult task; however, some progress in the geometrisation of the electromagnetic field has been made. This still leaves us with no connection between the theories of gravitational and nuclear forces; in fact, the whole relationship between general relativity and quantum theory is completely obscure. However, in recent years the concept of a potential has become an increasingly complex one in the field of nuclear physics, and there is growing feeling that much of this complexity arises from our viewing the problem from the wrong mental frame of reference. Who knows but that the clarification which we seek may lie in the analogous geometrisation of the nuclear field.

SYMMETRY AND THE CONSERVATION LAWS

We close this short account of modern physics with a brief excursion into a subject which is common to all branches of physics and whose principles were well established in classical physics. However, it is a subject whose great beauty has only been completely exposed with the development of modern physics.

We were introduced in Chapter 1 to two conservation laws: those of energy and linear momentum. Throughout the book other conservation laws have been implicitly assumed, e.g. conservation of charge and conservation of angular momentum. We might reasonably ask ourselves, what is the origin of these conservation laws, if they have an origin, and if not, are they simple accidents of nature?

If a system has any symmetry, then there will exist a set of transformations which preserve that symmetry, e.g. a sphere has a certain symmetry which is preserved by any rotation about its centre. In this context, a system could be some physical process, and one would expect this at least to be invariant with respect to simple transformations of the reference frame such as the Lorentz transformation (3.23). Associated with these symmetry transformations there will be some physical quantity which generates them, for example, linear spatial transformations are generated by linear momentum and spatial rotations by angular momentum. Thus it is clear that in such systems the quantity generating the symmetry transformations must be conserved, for, if it was not, we could use its variation to detect various orientations of the system, which, since we have said that it is symmetric, we are unable to do. Thus we see that the conservation of momentum is nothing but a statement of the fact that the universe appears to be symmetric with respect to spatial displacements, or, in other words, the laws of physics as formulated in London are the same as those formulated in New York. Then, when we remember that we do not live in a three-dimensional world but a four-dimensional space-time continuum and that in this world the fourth component of the momentum vector is the energy, we see that the conservation of energy merely gives expression to the observation that the universe is invariant with respect to a displacement of time, or the fact that the laws of physics are the same today as they were yesterday. Of course, the conservation of mass has, since the advent of relativity, become absorbed in the conservation of energy. It is because the gravitational potential depends only on the magnitude of the radius from its source and not on the orientation of that radius, i.e. the gravitational potential has spherical symmetry, that the Earth's angular momentum in its orbit is constant and that, as a result, the Earth has neither crashed into the Sun nor flown off into outer space.

We can go on looking for more and more symmetries, and with each one that we find we know that there must be some quantity which is conserved. The physical meaning of some of these symmetries becomes somewhat obscure; thus because the electromagnetic potential is invariant with respect to gauge transformations (a purely mathematical operation), electric charge is conserved. Other symmetries retain a clear physical significance; thus the fact that nuclear forces are symmetric with respect to mirror-like reflections gives rise to quantity, called the *parity*, which is conserved in any nuclear process. It is through listing the symmetry properties of the nuclear potential that we have been able to produce the classification of the baryons, mentioned in p. 101. However, many of these symmetries are not perfect but are obeyed only approximately. For example, if the symmetry of the nuclear potential was complete, all baryons would have the same mass; this they most certainly do not. Thus the symmetry must be broken slightly, but how and why?

Lastly, while it appears that the laws of physics have had a space-like and time-like symmetry during man's short study of them, will it always be so, and has it always been so in the past? Is it true everywhere in the universe or only in the proximity of the solar system? These and many questions like them are unlikely to find speedy solutions. They represent many of the most exciting, baffling and frustrating intellectual exercises that man has ever set himself, and no doubt they will continue to fascinate the human race as long as it is curious about its physical surroundings.

Appendix A

We have throughout this volume attempted to keep the use of mathematical techniques to a minimum. However, in order to keep the size of the book within limits, it has been necessary to use some mathematical symbols and nomenclature which may not be familiar to the reader. In this appendix we seek to define, without rigour, the terms we have used.

VECTOR ALGEBRA

We shall restrict ourselves to a simple Cartesian reference frame formed by three mutually perpendicular axes.

A vector **v** is defined by its three components (v_1, v_2, v_3) along three axes. The magnitude v of the vector is simply $(v_1^2+v_2^2+v_3^2)^{\frac{1}{2}}$ by Pythagoras's Theorem.

By adding two vectors, we mean that we form a vector whose components are the sum of the components of the original vectors, i.e. if $\mathbf{u}=\mathbf{v}+\mathbf{w}$ then **u** is the vector

$$(v_1+w_1,\ v_2+w_2,\ v_3+w_3).$$

To multiply a vector **v** by a scalar s we form a vector parallel to **v** but of length sv, i.e. if $\mathbf{u}=s\mathbf{v}$ then **u** is the vector (sv_1, sv_2, sv_3).

There are two forms of product of the vectors **v** and **w**. Firstly, there is the scalar product written as $\mathbf{v}\,.\,\mathbf{w}$, which is the scalar $v_1w_1+v_2w_2+v_3w_3$ or simply $vw\cos\theta$ where θ is the angle between **v** and **w**. Secondly, there is the vector product written as $\mathbf{v}\times\mathbf{w}$ which is a vector* perpendicular to both **v** and **w** and of magnitude $vw\sin\theta$ or simply

$$(v_2w_3-w_2v_3,\ v_3w_1-w_3v_1,\ v_1w_2-w_1v_2).$$

---

* We draw no distinction in this short appendix between vector and pseudo-vector quantities.

Vector quantities clearly depend on the reference frame in which they are measured. The same is not true of scalar quantities which are invariant with respect to the frame of reference.

DIFFERENTIAL OPERATORS AND VECTOR ANALYSIS

It has been assumed that the reader is familiar with elementary differential calculus as applied to simple scalar functions, i.e. if $f(x)$ is a scalar function of the variable x, then $\delta f/\delta x$ represents the rate of change of f with respect to x and $\delta^2 f/\delta x^2$ represents the rate of change of the rate of change of f with respect to x, etc. For example, if a particle is at the point s on the x-axis at the time t, then $\delta s/\delta t$ is the x-component of the velocity of the particle and $\delta^2 s/\delta t^2$ is the x-component of its acceleration.

We introduce the vector differential operator

$$\nabla=\left(\frac{\delta}{\delta x},\frac{\delta}{\delta y},\frac{\delta}{\delta z}\right).$$

When we multiply a scalar s by this operator we obtain a vector called the gradient of s, or simply grad s, which is the vector $\left(\frac{\delta s}{\delta x},\frac{\delta s}{\delta y},\frac{\delta s}{\delta z}\right)$. The scalar product of ∇ with a vector **v** is called the divergence of the vector, or simply div **v**, which is $\left(\frac{\delta v_1}{\delta x}+\frac{\delta v_2}{\delta y}+\frac{\delta v_3}{\delta z}\right)$. The vector product of ∇ with the vector **v** is called curl **v** and is the vector

$$\left(\frac{\delta v_3}{\delta y}-\frac{\delta v_2}{\delta z},\ \frac{\delta v_1}{\delta z}-\frac{\delta v_3}{\delta x},\ \frac{\delta v_2}{\delta x}-\frac{\delta v_1}{\delta y}\right).$$

The scalar product of ∇ with itself is written as ∇^2 and it is a scalar quantity, i.e. acting on the scalar s we obtain the scalar $\frac{\delta^2 s}{\delta x^2}+\frac{\delta^2 s}{\delta y^2}+\frac{\delta^2 s}{\delta z^2}$, acting on the vector **v** we obtain the vector

$$\left(\frac{\delta^2 v_1}{\delta x^2}+\frac{\delta^2 v_1}{\delta y^2}+\frac{\delta^2 v_1}{\delta z^2},\ \frac{\delta^2 v_2}{\delta x^2}+\frac{\delta^2 v_2}{\delta y^2}+\frac{\delta^2 v_2}{\delta z^2},\ \frac{\delta^2 v_3}{\delta x^2}+\frac{\delta^2 v_3}{\delta y^2}+\frac{\delta^2 v_3}{\delta z^2}\right).$$

In the case of the electrostatic and gravitational fields, the field strength for a charge of mass at the origin is radial and proportional to $1/r^2$, i.e. the vector field is proportional to the vector $\left(\frac{x}{r^3}, \frac{y}{r^3}, \frac{z}{r^3}\right)$ where $r=(x^2+y^2+z^2)^{\frac{1}{2}}$. The potential is thus proportional to $\int \frac{xdx}{r^3}$, equation (1.3), i.e. **v** is proportional to $-\frac{1}{r}$. The gradiant of $-\frac{1}{r}$ is the vector $\left(\frac{x}{r^3}, \frac{y}{r^3}, \frac{z}{r^3},\right)$ as the reader may easily verify.

Many relationships between the differential operators grad, div and curl can be obtained; we shall only prove the single result used on p. 24, namely curl (curl **E**) = grad (div **E**) $-\nabla^2$**E**. Each side of this equation is a vector, and we shall establish the equality for the first component of this vector, the result for the other two components follows in an identical fashion.

The first component of the right-hand side is

$$\frac{\delta}{\delta x}\left(\frac{\delta E_1}{\delta x}+\frac{\delta E_2}{\delta y}+\frac{\delta E_3}{\delta z}\right)-\left(\frac{\delta^2 E_1}{\delta x^2}+\frac{\delta^2 E_2}{\delta y^2}+\frac{\delta^2 E_1}{\delta z^2}\right)$$

$$=\frac{\delta^2 E_2}{\delta x \delta y}+\frac{\delta^2 E_3}{\delta x \delta z}-\frac{\delta^2 E_1}{\delta y^2}-\frac{\delta^2 E_1}{\delta z^2}. \qquad \text{(A. 1)}$$

The first component of the left-hand side is

$$\frac{\delta}{\delta y}(\text{curl } \mathbf{E})_3-\frac{\delta}{\delta z}(\text{curl } \mathbf{E})_2=\frac{\delta}{\delta y}\left(\frac{\delta E_2}{\delta x}-\frac{\delta E_1}{\delta y}\right)+\frac{\delta}{\delta z}\left(\frac{\delta E_3}{\delta x}-\frac{\delta E_1}{\delta z}\right)$$

$$=\frac{\delta^2 E_2}{\delta y \delta x}+\frac{\delta^2 E_3}{\delta z \delta x}-\frac{\delta^2 E_1}{\delta y^2}-\frac{\delta^2 E_1}{\delta z^2}. \qquad \text{(A. 2)}$$

Since the order of differentiation is interchangeable, we see that (A. 1) and (A. 2) are indeed identical.

Relativistic forms

In a relativistic theory it is convenient to work in a four-dimensional frame, since now there is no absolute time scale.

All vectors now have four components, three of them space-like and one time-like. The fourth component is always imaginary (in the mathematical sense that it is proportional to i, the square root of -1) if a system is chosen in which the space-like components are real. We content ourselves by simply giving a few examples of four-vectors.

1. The simple displacement vector becomes (x, y, z, ict) where c is the velocity of light and t is the time.
2. The momentum vector **p** becomes $(p_1, p_2, p_3, i\mathrm{E}/c)$ where E is the energy.
3. The current vector **j** becomes $(j_1, j_2, j_3, ic\rho)$ where ρ is the charge density.
4. The relativistic analogue of the operator ∇ is written as $\Box$ and is the vector operator $\left(\frac{\delta}{\delta x}, \frac{\delta}{\delta y}, \frac{\delta}{\delta z}, -\frac{i}{c}\frac{\delta}{\delta t}\right)$.

The fourth component of the vector is chosen so that the scalar product of the vector with itself will be a scalar in the four-dimensional reference frame.

In this formalism the equations of electromagnetism take on a new simplicity, e.g. equation (2.4) is simply $\Box . \mathbf{j}=0$, and the wave equation (2.11) becomes $\Box^2\mathbf{E}=0$, etc.

Appendix B

In this appendix we attempt a brief introduction to the statistics of large numbers of particles. No attempt is made at either mathematical rigour or completeness, since this would take us far beyond the scope of this small volume.

AVERAGES AND DISTRIBUTION FUNCTIONS

Suppose we make a set of N simple observations and obtain the results $x_1, x_2, \ldots, x_N$, then to obtain the average result we simply add them together and divide by the total number of observations, i.e.

$$\langle x\rangle=\frac{1}{N}\sum_{i=1}^{N} x_i. \tag{B. 1}$$

If the observation x_1 occurred n_1 times, the observation x_2 occurred n_2 times, etc., then we could write

$$\langle x\rangle=\frac{1}{N}\,(n_1x_1+n_2x_2+\ldots+n_Nx_N) \tag{B. 2}$$

$\frac{n_1}{N}$ is a measure of the probability of the result of the observation being x_1, etc. If we denote the probability of the result being x_i by fi, then (B. 2) becomes

$$\langle x\rangle=\frac{\sum\limits_{i=1}^{N} fi\ x_i}{\sum\limits_{i=1}^{N} fi} \tag{B. 3}$$

where N is the number of different results of the experiment. If there are very many results which have a continuous distribution of values, then we replace the discrete summations in (B. 3) by integrals

$$\langle x\rangle=\frac{\int_{-\infty}^{\infty} f(x)x\,dx}{\int_{-\infty}^{\infty} f(x)\,dx}. \tag{B. 4}$$

Since multiplying $f(x)$ by a constant does not change $\langle x\rangle$, we can choose $f(x)$ such that $\int_{-\infty}^{\infty} f(x)dx=1$. Then $f(x)$ is called the probability distribution function, and the probability that a result will lie in the range x to $x+dx$ is $P(x, dx)=f(x)dx$. The condition that the integral of $f(x)$ is unity makes sense, since it simply asserts that the sum of all possibilities is unity, i.e. we must obtain some result for our observation.

Suppose that we make a number of simultaneous observations of the quantities x, y and z which have distribution functions f_1, f_2 and f_3 respectively. Then, the probability that the results will lie in the ranges x to $x+dx$, y to $y+dy$ and z to $z+dz$ is simply the product of the individual probabilities, i.e. $P(x, dx, y, dy, z, dz)=f_1(x)\,f_2\,(y)\,f_3\,(z)\;dz\;dy\;dz$.

BOLTZMANN STATISTICS

Let us consider the non-interacting gas described on p. 14 and calculate the distribution of velocities.

From the symmetry of the problem it is clear that the distribution function cannot depend on the direction of the velocity **u** but only on its magnitude, $u=(u_1^2+u_2^2+u_3^2)^{\frac{1}{2}}$; also, the distribution functions for each component must be the same. This means

$$f(u_1)\,f(u_1)\,f(u_3)=f(u). \tag{B. 5}$$

Differentiating (B. 5) by u_1 and dividing by $f(u)$, we have that

$$\frac{\delta f(u_1)}{\delta u_1}\Big/ u_1 f(u_1)=\frac{\delta f(u)}{\delta u}\Big/ uf(u),$$

and, similarly,

$$\frac{\delta f(u_2)}{\delta u_2}\Big/ u_2 f(u_2)=\frac{\delta f(u_3)}{\delta u_3}\Big/ u_3 f(u_3)=\frac{\delta f(u)}{\delta u}\Big/ u f(u). \qquad \text{(B. 6)}$$

But each of these expressions is, in turn, only a function of the variables u_1, u_2, u_3 and u respectively. Hence, for them always to be equal, we require that they must all be equal to a constant which we shall choose to write in the form $-1/\alpha^2$. Thus

$$\frac{\delta f(u)}{\delta u}=-\frac{1}{\alpha^2}uf(u) \qquad \text{(B. 7)}$$

or

$$f(u)=A\exp(-u^2/\alpha^2)$$

In order that $\int_{-\infty}^{\infty} f(u)du=1$ then $A=1/\alpha\sqrt{\pi}$

The average kinetic energy $\langle E\rangle$ per molecule is

$$\langle E\rangle=\frac{4}{\alpha^3\sqrt{\pi}}\int(\tfrac{1}{2}mu^2)\exp(-u^2/\alpha^2)u^2\,du=\tfrac{1}{2}m(\tfrac{3}{2}\alpha^2). \qquad \text{(B. 8)}$$

However, on p. 16 we expressed the average kinetic energy per molecule in terms of the temperature T as $\frac{3}{2}kT$. Thus $\alpha^2=2kT/m$ and the Boltzmann distribution function is proportional to $\exp(-\frac{1}{2}mu^2/kT)$. The result can be generalised to particles which have potential as well as kinetic energy, with the result that the energy distribution function $f(E)$ is proportional to $\exp(-E/kT)$.

Suppose that the energy contains a quadratic term, i.e. $E=E_0+q^2$, then the average energy is

$$\begin{aligned}\langle E\rangle&=\textstyle\int(E_0+q^2)\exp[-(E_0+q^2)/kT]\,dq/\int\exp[-(E_0+q^2)/kT]\,dq\\&=E_0+\textstyle\int q^2\exp(-q^2/kT)\,dq/\int\exp(-q^2/kT)\,dq\\&=E_0+\tfrac{1}{2}kT. \qquad \text{(B. 9)}\end{aligned}$$

That is, on average the quadratic term contributes $\frac{1}{2}kT$ to the total energy.

QUANTUM STATISTICS

Consider an assembly of one-dimensional oscillators, then, according to the result (B. 9), they will have an average energy

of kT (i.e. $\frac{1}{2}kT$ from the kinetic energy and $\frac{1}{2}kT$ from the potential energy). This was the result assumed by Rayleigh and Jeans in arriving at equation (4.4). Now suppose that the oscillators are quantised and can only have energies $nh\nu$ where n is an integer. Then the average energy is

$$\langle E\rangle=\sum_{n=0}^{\infty}nh\nu\exp(-nh\nu/kT)\Big/\sum_{n=0}^{\infty}\exp(-nh\nu/kT). \qquad \text{(B. 10)}$$

The numerator is
$h\nu\exp(-h\nu/kT)[1+2\exp(-h\nu/kT)+3\exp(-zh\nu/kT+\ldots]$ and is simply $h\nu\exp(-h\nu/kT)[1-\exp(-h\nu/kT)]^{-2}$ by the binomial theorem. The denominator is the sum of a simple geometric series and has the value $[1-\exp(-h\nu/kT)]^{-1}$. Thus

$$\langle E\rangle=h\nu[\exp(h\nu/kT)-1]^{-1}. \qquad \text{(B. 11)}$$

Using this result with $\nu=c/\lambda$, Planck obtained the equation (4.6) and thus solved the problem of black-body radiation. Also, if we consider a solid to be made up of $3N$ such oscillators, then the specific heat is

$$\begin{aligned}&\frac{\delta}{\delta T}\left(3Nh\nu\left[\exp\left(\frac{h\nu}{kT}\right)-1\right]^{-1}\right)\\&=3Nk\left(\frac{h\nu^2}{kT}\right)\exp\left(\frac{h\nu}{kT}\right)\left[\exp\left(\frac{h\nu}{kT}\right)-1\right]^{-2}. \qquad \text{(B. 12)}\end{aligned}$$

i.e. Einstein's result referred to on p. 79.

The distribution function $[\exp(h\nu/kT)-1]^{-1}$ can be generalised to any set of quantised energy levels $E_1, E_2, \ldots$, in which case it becomes $[\exp([E_k-\mu]/kT)-1]^{-1}$ where μ, the chemical potential, is a reference energy chosen to give

$$\sum_K[\exp([E_k-\mu]/kT)-1]^{-1}=N,$$

the total number of particles in the gas, since now we look on the distribution function as the probability that a particle will have a given energy, and if we sum over all possible energies we must cover all particles. This is the distribution function governing bosons for which the energy levels are quantised, but there is no limit on how many particles can enter any state.

For fermions, the number of particles in any state is either 0 or 1, and so (B. 10) becomes simply

$$\langle E\rangle = h\nu \exp(-h\nu/kT)[1+\exp(-h\nu/kT)]^{-1}. \qquad \text{(B. 13)}$$

Similarly, this distribution function can be generalised and becomes $[\exp([E_k-\mu]/kT)+1]^{-1}$. We see that for E_k much greater than μ the probability of the state being occupied is very small, for $E_k=\mu$ the probability is $\frac{1}{2}$ and for E_k smaller than μ the probability tends to 1 (Fig. 32). Thus even at very low temperatures fermion systems have a finite energy structure. It is this feature which gives our world its physical shape and prevents it from becoming supercondensed. For example, the volume of an atomic nucleus is approximately 10^{-39} cm^3, but the volume of an atom is approximately 10^{-24} cm^3 even although the actual volume of the electrons is negligible.

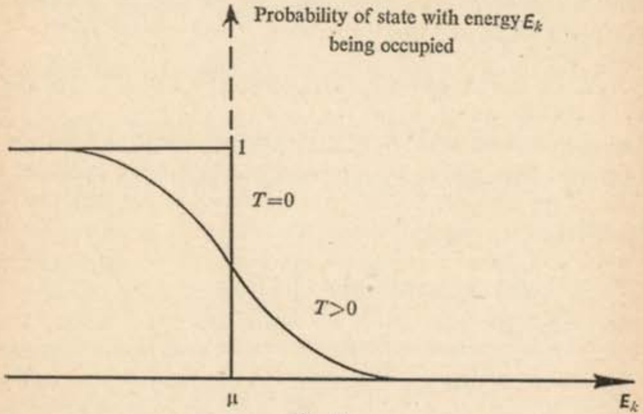


Fig. 32.

Bibliography

1. CLASSICAL MECHANICS AND THE STRUCTURE OF MATTER

1963. *Feynman lectures on physics*, Vol. I. Addison-Wesley, Reading, Mass.
BORN, MAX. 1962. *Atomic physics*, Chap. 1. Blackie, London.
* BORN, MAX. 1962. *Einstein's theory of relativity*, Chaps. 1-3. Dover Publications, New York.

2. ELECTROMAGNETISM, LIGHT AND WAVES

1963, 1964. *Feynman lectures on physics*, Vols. I and II. Addison-Wesley, Reading, Mass.
* BORN, MAX. 1962. *Einstein's theory of relativity*, Chaps. 4-5. Dover Publications, New York.

3. RELATIVITY

* BORN, MAX. 1962. *Einstein's theory of relativity*, Chap. 6. Dover publications, New York.
RICHTMYER, KENNARD AND LAURITSEN. 1965. *Introduction to modern physics*, Chap. 2. McGraw-Hill, New York.
EISBERG, R. M. 1961. *Fundamentals of modern physics*, Chap. 1. Wiley, New York.
* EINSTEIN, A. 1954. *Relativity*, Part I. Methuen, London.

4. THE INTERACTION OF THE ELECTROMAGNETIC FIELD WITH MATTER

EISBERG, R. M. 1961. *Fundamentals of modern physics*, Chaps. 2-4. Wiley, New York.
RICHTMYER, KENNARD AND LAURITSEN. 1965. *Introduction to modern physics*, Chap. 4. McGraw-Hill, New York.
* SLATER, J. C. 1955. *Modern physics*, Chaps. 1 and 3. McGraw-Hill, New York.

5. THE OLD QUANTUM THEORY

RICHTMYER, KENNARD AND LAURITSEN. 1965. *Introduction to modern physics*, Chap. 5. McGraw-Hill, New York.
* SLATER, J. C. 1955. *Modern physics*, Chaps. 3-5. McGraw-Hill, New York.

6. QUANTUM MECHANICS

EISBERG, R. M. 1961. *Fundamentals of modern physics*, Chaps. 6-12. Wiley, New York.

1965. *Feynman lectures on physics*, Vol. III. Addison-Wesley, Reading, Mass.

* HOFFMANN, B. 1959. *The strange story of the quantum*. Penguin Books, London.

7. POT-POURRI

* YANG, C. N. 1962. *Elementary particles*. Princeton Univ. Press, Princeton, N.J.

RICHTMYER, KENNARD AND LAURITSEN. 1965. *Introduction to modern physics*, Chaps. 10-11. McGraw-Hill, New York.

* COOK, C. SHARP. 1964. *Structure of atomic nuclei*. Van Nostrand, New York.

* BORN, MAX. 1962. *Einstein's theory of relativity*, Chap. 7. Dover Publications, New York.

* EINSTEIN, A. 1954. *Relativity*, Part II. Methuen, London.

APPENDIX A

1963, 1964. *Feynman lectures on physics*, Vol. I, Chap. 11; Vol. II, Chap. 2. Addison-Wesley, Reading, Mass.

APPENDIX B

1963. *Feynman lectures on physics*, Vol. I, Chaps. 6, 39, 40. Addison-Wesley, Reading, Mass.

BORN, MAX. 1962. *Atomic physics*, Chap. 8. Blackie, London.

Books marked * contain only the minimum of mathematics.

Index

Absolute time scale, 46
Absorption wavelengths, 57
Absorptivity, 58
Alkali metals, 73, 92
α-particles, 67, 94, 102
α-rays, 66
Ampère, 23
Ampere unit, 13
Anderson, 99
Angle of incidence, 29
Angle of reflection, 29
Angular momentum, quantisation of, 69, 75, 90
Angular momentum quantum number, 92
Antinodes, 28
Antiparticles, 99
Atomic bomb, 103
Atomic reactors, 55, 103
Atomic weight, 12, 96, 97
Atomicity of matter, 11-14
Averages, 114-15
Avogadro's hypothesis, 12
Azimuthal quantum number, 76

Balmer series, 71
Baryons, 100, 109
Becquerel, 65
β-decay, 102
β-rays, 66
Black-body radiation, 58-61
Bohr, 75, 77
Bohr orbit, 89, 92
Bohr's atom, 69-71, 75, 76, 88
Boltzmann constant, 16
Boltzmann statistics, 17, 58, 115-116
Bose, 94
Bosons, 94, 117
Boyle's Law, 17
Brackett series, 71
Bunsen, 56

Carlisle, 13
Cartesian reference frame, 110
Cathode ray, 63
Chadwick, 96
Charge density, 22
Charles's Law, 17
Chemical potential, 117
Classical mechanics, 3-10
Compass, 22
Conduction of electricity, 18
Conservation laws, 6-8, 107-9
Constant and multiple proportions, laws of, 11
Constant volumes, law of, 11
Copernicus, 40
Corpuscular theory of light, 33, 35, 36
Correspondence Principle, 77-8
Cosmic rays, 98, 99
Coulomb unit, 13
Critical frequency, 65
Critical mass, 103
Curie, 65, 96

Dalton, 12
Davisson, 81, 82
De Broglie, 79
De Broglie relations, 81, 84, 85

Debye, 79
Differential operators, 111-13
Diffraction pattern, 86, 88
Dirac, 79, 99
Dispersion, 30-1
Distribution functions, 114-15
Dulong, 18, 78

Earth's magnetic field, 22
Ehrenfest, 68
Einstein, 46 ff.
Einstein's General Theory of Relativity, 55, 104-7
Einstein's Special Theory of Relativity, 46-55, 77
Electric charge, 10, 13, 20, 109
Electrolysis, 13
Electromagnetic field, 20-5
Electromagnetic spectrum, 38
Electron microscopes, 83
Electrons, 62, 63, 84, 97, 99
Electrostatics, 10
Elementary particles, 81-8, 96-101
Emissivity, 57
Ether, 34, 40, 45
  wind, 41, 43
Exclusion Principle, 88-95, 99

Faraday, 13, 23
Faraday Effect, 32
Fermi, 90, 94
Fermions, 94, 97
Fission, 101-3
Fitzgerald, 45
Fitzgerald-Lorentz contraction, 46, 49, 55
Fizeau, 36
Foucault, 36
Frank, 74
Fraunhofer dark lines, 56, 57
Fusion, 103-4

Galilean Transformation, 42, 48
Galileo, 40
γ-rays, 66
Gas pressure, 16
Gedanken experiment, 2, 104
Gell-Mann, 101
General Theory of Relativity, 9, 55, 104-7
Geodesics, 105
Gerlach, 90
Germer, 81, 82
Gilbert, 22
Goudsmit, 91
Gravitation, law of, 8

H-bomb, 103
Hallwachs, 61
Heisenberg, 79, 88
Heisenberg Uncertainty Principle, 81-8
Helium, 57
Hertz, 24, 61, 74
Huygens, 34
Hydrogen atom, 72, 88

Index of refraction, 30, 33, 36
Inert gases, 92
Interference, 29

Jeans, 16
Joliot, 96

Kepler, 40, 106
Kinetic theory of matter, 14-19
Kirchhoff, 56, 58

Lenard, 62, 63
Light
  diffraction of, 34
  nature of, 33-9
  polarisation of, 31-2, 33
  properties of, 29-32
  velocity of, 36-8, 47, 49
Line spectra, 56-8
Lodestone, 22
Lorentz, 45
Lorentz Force, 25
Lorentz Transformation, 48, 50
Lyman series, 71

Magnetic fields, 10, 22
Magnetic poles, 22
Mass-energy relationship, 55
Maxwell, 38
Maxwell's Displacement Current, 22
Maxwell's Equations, 24, 55
Mendeleeff, 11
Meson, 98
Michelson, 40
Michelson-Morley Experiment, 40-6
Millikan, 65
Mole, 12
Molecular weight, 12
Morley, 40
Muons, 98, 99

Ne'eman, 101
Negative energy states, 99
Neutrino, 97
Neutron decay, 97
Neutron flux, 103
Neutrons, 96, 97, 102
Newton, 31, 33, 41, 56
Newtonian relativity, 40-6
Newton's Second Law of Motion, 6
Newton's universal law of gravity, 8
Nicholson, 13
Nodes, 28
Nuclear fields, 10
Nuclear forces, 98, 101
Nucleons, 101
Nucleus, 68

Oersted, 23
Ohm's Law, 19
Ohnuki, 101
Old Quantum Theory, critique of, 79-80

Parity, 109
Particle accelerators, 55
Particles, *see* Elementary particles
Paschen series, 71
Pauli, 79 ff.
Pauli Exclusion Principle, 88-95, 99
Periodic Table of the Elements, 93
Petit, 18, 78
Pfund series, 72
Photo-electric effect, 61-8
Photons, 65, 86, 94
Physics, laws of, 2
Pions, 95, 98
Planck, 59 ff.
Planck's constant, 65 ff.
Planck's postulate, 60, 75
Polaroid, 32
Positron, 99
Powell, 98
Protons, 72, 97, 102
Prout, 96
Pythagoras's Theorem, 44, 47, 110

Quantum conditions, 75
Quantum number, 60, 76
Quantum state, 60, 95
Quantum statistics, 116-17

Radio waves, 38
Radioactivity, 66
Rainbow, 31
Rayleigh-Jeans Law, 58-9, 60
Red shift, 106
Reflection, 29
Refraction, 29
Relativistic forms, 112-13
Rest mass, 55
Riemannian geometry, 105
Romer, 36
Rutherford, 65 ff.
Rutherford's nuclear atom, 65-8
Rydberg constant, 72

Salam, 101

Schroedinger, 79, 88
Selection rules, 77-8
Simple harmonic wave, 26-7
Snell's Law, 29
Special Theory of Relativity, 46-55, 77
Specific heat, 17
Specific heats of solids, 78-9
Spectral analysis, 56
Spectral lines, 71-5
Spectrum, 31
Spin, 88-95
Standing wave, 27-8
Stefan-Boltzmann constant, 57, 58
Stefan's Law, 57, 58
Stellar energies, 104
Stern, 90
Stoletow, 61
Structure (of atoms), 88-95
Superposition, Principle of, 20, 27
Symmetry, 107-9

Thermal energy, 16
Thermal radiation, 58
Thermodynamic equilibrium, 58
Thin films, 35
Thomas, 90
Thomson, 62, 66
Thorium, 66
Time dilation, 50, 55

Uhlenbeck, 91
Ultra-violet catastrophe, 59
Uncertainty Principle, 81-8
Unified field theory, 107
Uranium, 66, 102

Valency electrons, 94
Vector algebra, 110-11
Vector analysis, 6, 111-13

Wave front, 48
Wave packet, 85, 86, 87
Wave theory of light, 34
Wave velocity, 86
Wavelength, 27
Waves, 25-9, 81-8
Wien's Law, 58, 59, 60
Wilson-Sommerfeld quantisation rules, 75-7

X-ray crystallography, 82
X-ray diffraction techniques, 13
X-rays, 82

Young, 34, 35
Yukawa, 98

Zeeman effect, 90

Other titles in this series
from Oliver and Boyd
are described on the following pages

CSP 1

BASIC ASTRONOMY

Patrick Moore

It is an astonishing fact that although Aristarchus of Samos was seriously suggesting that the Earth moved round the Sun, his ideas did not prevail over the Ptolemaic System and it was left to Copernicus to spark off the great astronomical awakening of the Renaissance.

In this book Patrick Moore traces briefly the history of astronomy, developing his narrative into a full account of Earth and sky, the Sun, the Solar System, the stars, the galaxies, and quasi-stellar objects, or quasars. Such knowledge does not come from theorising alone, but from innumerable observations through the agencies of the naked eye, optical and radio telescopes, and the spectroscope. The author describes these instruments together with the principles of their use.

It is also the business of astronomy to speculate about the origins and evolution of the universe. We are therefore given a glimpse of the various theories that have been proposed from the Abbé Lemaitre's notion of the primeval atom in the early 1920s to present-day speculations. We are left with the thought that it is surely the height of conceit to suppose that our particular Sun is unique in being attended by a life-bearing planet.

Patrick Moore is well known to the general public through his other books and his BBC television programme *The Sky at Night*. He is Director of the Armagh Planetarium in Northern Ireland.

CSP 2

LIFE, MIND AND GALAXIES

V. A. Firsoff

First and foremost an astronomer, Val Axel Firsoff also has a vast knowledge of experimental facts in biology, chemistry, and biochemistry. With a fine intuition on the nature of mind in relation to the rest of physical reality he is able to look on the natural world with a breadth of vision that is rare nowadays.

In this book he reflects on the origins of life, examines its chemical basis, and illustrates the possibility of alternative biochemistries for life elsewhere in the universe. The warp and weft of his cosmological fabric are living things and the universe, both of which are shown to evolve towards higher patterns of organisation. In each case experimental evidence supports the argument.

Perhaps the most controversial part of this work is the contention that mind is an essential part of physical reality, which leads the author to discuss the possible relationship between subatomic physics, consciousness, and free will.

V. A. Firsoff is the author of numerous scientific papers and has lectured widely to university and other societies on the subject of the universe. This is his twentieth book.

CSP 3

REPRODUCTION AND MAN

Richard J. Harrison

Control of reproduction demands a full knowledge of reproductive patterns and organs, and of the factors that influence them. It demands also that full consideration be given to the possible dangers of short- and long-term interference in reproductive events for the purpose of fertility control.

Professor Harrison therefore describes the biological characteristics of reproduction in man and compares them with those exhibited by other mammals. He reviews in simple terms the essential phases of growth, activity, and control of reproductive organs and processes, emphasising those features where increased scientific knowledge might allow control and improvement of fertility.

Richard J. Harrison is Professor of Anatomy at the London Hospital Medical College and Fullerian Professor of Physiology at the Royal Institution. He has written many research papers on various aspects of reproduction and embryology, and he is particularly interested in the curious reproductive patterns of marine mammals such as seals and dolphins.

CSP 4

MOLECULAR BIOPHYSICS

D. Chapman/R. B. Leslie

Molecular biophysics is concerned with the properties of biological systems; it is directed towards an understanding of those systems in terms of the physical and chemical properties of the constituent atoms and molecules and uses physical techniques as its main tool for investigation. Molecular biophysics therefore is the meeting point of biology, chemistry and physics—an alliance that has produced some of the most striking discoveries of the twentieth century.

This book describes what molecular biophysics is about. It ranges from a general introduction to the subject to a detailed examination of selected topics such as vision and the transmission of nerve impulses. Further because the conversion of energy from one form to another in biological systems involves electron migration we are shown how molecular biophysics leads into the realm of subatomic events and therefore to quantum biophysics. Indeed it is part of biophysics to question the nature of the continuum of understanding between subatomic events and living things.

Dennis Chapman and **Robert Leslie** both work at the Unilever Research Laboratory, Welwyn, where Dr Chapman is Head of the Molecular Biophysics Unit and also Head of the General Research Division.

CSP 5

NOISE AND SOCIETY

Michael Rodda

Everyone agrees that noise is annoying but we are not entirely sure why. Undoubtedly when we speak of noise we are referring to something unwanted. Efforts have therefore been made to restrict the amount of noise in society and various techniques have been developed to assess sound so that the measure has a psychological meaning. The scientific measurement of noise is not difficult but its control is complicated by the fact that instruments do not react to noise in the way a human being does. The same measured intensity of sound may be pleasant or objectionable depending on whether it is produced by a symphony orchestra or a pneumatic road drill. In turn, even the orchestral sound may appear as unwanted noise to a non-musical person.

Dr Rodda first defines noise and in subsequent chapters deals with legal aspects, measurement, occupational deafness, and the effects of noise on speech efficiency, mental health and physiological disturbance. He considers also the many sources of noise in our environment and deals with the general questions of its control in the community. Throughout the book the author maintains the psychological as well as the physical approach and presents a personal interpretation of current knowledge of the subject.

Michael Rodda is Medical Research Council Lecturer in Experimental Psychology in the Department of Audiology and Education of the Deaf at Manchester University.

CSP 6

THE IDEAS OF PHYSICS

Ernest H. Hutten

While deliberate experimentation is characteristic of modern science, it is true to say, with A. N. Whitehead, that 'all science as it grows towards perfection becomes mathematical in its ideas'. It is important not to misinterpret this view, because the main import of mathematics in the physical sciences is that it provides a universal system of symbols rather than a means for quantitative judgement. The abstract character of mathematics enables the physicist to push his theories beyond the limits of ordinary language into the realm of atoms and galaxies. Confirmation of a hypothesis, however, is not a matter of mere numerical agreement. It depends on the interpretation of the mathematical formalism, an interpretation which in the last resort expresses rational human behaviour.

Dr Hutten develops this theme in relation to the ideas of physics, and he constantly illustrates the fact that science—and in particular, physics—is a creative activity, that the external world and the inner world of man are mutually dependent, and that knowledge is the result of interactions between them. He shows us how physics exemplifies the way in which a new theory is built upon an old one, and how the old theory is in turn corrected by the new one. This is the correspondence and self-correction which characterises the 'open' mode of scientific thinking. This openness and freedom makes science the most advanced kind of thought mankind has so far developed.

Ernest H. Hutten is Reader in Physics, Royal Holloway College, London University.

CSP 7

ATTRACTION AND REPULSION

Malcolm McCaig

Although there are many situations in which permanent magnets have a mechanical application, little enough attention has been given to the subject. In this book Dr McCaig describes the manner in which permanent magnets exert force and transmit power. He illustrates the features of permanent magnets used in simple attracting, holding, and lifting devices, and those that can be switched on or off, such as permanent magnetic chucks. Other applications exploit the property of repulsion, as in the sheet floater, frictionless bearings, dentures, and the support of vehicles—suggesting the intriguing problem of levitation without mechanical contact. Among the many uses of permanent magnets are also those that relate to electric motors, generators, magnetic drives, brakes, and contactless gears. Even to the engineer familiar with electromagnetic devices much of this account will be novel.

Malcolm McCaig is Assistant Director of Research at the Central Research Laboratory of the Permanent Magnet Research Association in Sheffield.

CSP 8

AUTOMATION: ITS ANATOMY AND PHYSIOLOGY

John Rose

Beginning with a review of the concept of automation in its various settings, Dr Rose turns his attention to a detailed analysis of the three basic features of automated systems—computer, feedback, and automatic control.

Tracing back the origin of mechanisation to the first Industrial Revolution, the author makes the distinction between the tools of the period—machines working under detailed human supervision—and the tools of the second Industrial Revolution—computers—which take over from man his control function.

Self-controlling machines, are therefore the subject of automation. Actuated by computer, the complete automaton can simulate the workings of the human mind in the execution of certain tasks. Using feedback action, made effective in the technology of automatic control, the automaton is able to make decisions by modifying its own programme of instructions so as to achieve a given objective.

It goes without saying that the implications of automation are tremendous. The reader who wishes to learn about the present applications of automation, its economic and social consequences may turn to CSP 9, also by Dr Rose, entitled *Automation: its uses and consequences.*

John Rose is Principal of the College of Technology and Design in Blackburn. With his essay on *The problem of technological barriers* he won first prize in the 1966 international competition organised by Shell Chemicals Ltd and the British Association for the Advancement of Science. He is especially interested in computer science education.

CSP 9

AUTOMATION: ITS USES AND CONSEQUENCES

John Rose

Although advanced industrial societies are already committed to automation, where machines and processes are controlled by computer instead of by man, effective planning for automation on a large scale demands a more widespread knowledge of the likely consequences than there is at present.

Dr Rose considers automation in relation to its uses and describes its mode of operation in industry, commerce, public administration, banking and so on. This points already to some of the economic, labour, management, and social problems of automation with which the author deals in turn. Truly, automation affects people in every place of employment, not only on the workshop floor but also in the boardroom. New skills have to be acquired and greater mental dexterity must be developed. This is precisely the problem of automation, because 'one of the greatest pains of human nature is the pain of a new idea'.

The reader wishing to learn more about the scientific ideas behind automation may turn to CSP 8, also by Dr Rose, entitled *Automation: its anatomy and physiology.*

John Rose is Principal of the College of Technology and Design in Blackburn. With his essay on *The problem of technological barriers* he won first prize in the 1966 international competition organised by Shell Chemicals Ltd and the British Association of the Advancement of Science. He is especially interested in computer science education.

The Basis of Modern Physics

John M Irvine

Modern physics is concerned with those branches of physics which, for their theoretical interpretations, depend upon ideas developed during the twentieth century. These ideas are necessary to explain observations made beyond the normal limits of the macroscopic world. For example, when we consider systems moving with speeds approaching that of light, we require the theory of relativity in place of our usual concepts of space and time; when we discuss the structure of the atom we need quantum theory instead of the mechanics of Newton.

In this elementary introduction to modern physics, Dr Irvine sets the stage by giving us an account of the results of classical physics—from the laws of classical mechanics to Maxwell's electromagnetic theory. He then considers Einstein's special theory of relativity, together with some of its startling predictions.

While many simple accounts of relativity are available, simple treatments of quantum theory are less common, and this has led the author to deal with the latter at greater length than is usual in books of this size. He therefore includes three chapters concerned respectively with the problem of the interaction of the electromagnetic field with matter, the old quantum theory, and modern quantum mechanics. He concludes with a collection of problems, some partially solved and some very far from solution, which face modern physicists.

John Maxwell Irvine is Lecturer in Theoretical Physics at Manchester University.

Contemporary Science Paperbacks are lucid, compact and authoritative. Through them, scientific and technological ideas and achievements are made readily comprehensible.

Recommended Overseas Prices

Australia $A1.15	Canada $1.65	Italy L850
New Zealand 9s. 6d.	South Africa R0.93c.	

John M Irvine

The Basis of Modern Physics

12